

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 112.61 Seconds

(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-2

Perfect score: 1 DIQLTQSPSSILAVSAGENVY.....CHQTLSSWTFGGTKLEIKR 113

Sequence: 1 DIQLTQSPSSILAVSAGENVY.....CHQTLSSWTFGGTKLEIKR 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.23:*
1: SP_Archaea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Pnige:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_Virus:*
16: SP_Bacteriap:*
17: SP_Archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418.5	70.5	107	11	Q9ER29
2	392	66.0	109	4	Q9UL78
3	386	65.0	107	4	Q96SA9
4	383.5	64.6	108	4	Q9UL77
5	375.5	63.2	108	4	Q9UL70
6	373	62.8	107	4	Q9UL81
7	371	62.5	107	4	Q99M37
8	370	62.3	109	4	Q9UL85
9	368	62.0	239	4	Q8NEK0
10	367	61.8	109	4	Q9UL86
11	366	61.6	238	11	Q8VC56
12	362.5	61.0	108	4	Q9UL83
13	360.5	60.7	99	11	Q9UL74
14	360.5	60.7	108	4	Q9UL79
15	359.5	60.5	234	4	Q8NEK1
16	359	60.4	239	11	Q8VC55

17	356.5	60.0	111	11	Q920E9	Q920E9 mus musculus
18	355.5	59.8	114	4	Q9UL80	Q9UL80 homo sapien
19	353.5	59.5	233	11	Q91WS9	Q91WS9 mus musculus
20	350	58.9	239	11	Q8K0F8	Q8K0F8 mus musculus
21	349.5	58.8	108	11	Q8VIC0	Q8VIC0 mus musculus
22	348	58.6	239	4	Q8RCDO	Q8RCDO homo sapien
23	346.5	58.3	109	11	Q920B6	Q920B6 mus musculus
24	342.5	57.7	214	11	Q9RIAS	Q9RIAS mus musculus
25	339	57.1	104	11	Q9JL82	Q9JL82 mus musculus
26	336.5	56.6	298	11	Q9QYF0	Q9QYF0 mus musculus
27	333.5	56.1	234	11	Q8RD60	Q8RD60 mus musculus
28	333	56.1	134	11	Q8VDD0	Q8VDD0 mus musculus
29	332	55.9	112	11	Q8KIF3	Q8KIF3 mus musculus
30	330	55.6	114	11	Q8KIF1	Q8KIF1 mus musculus
31	322.5	54.3	107	11	Q9JL84	Q9JL84 mus musculus
32	319.5	53.8	103	11	Q9JL80	Q9JL80 mus musculus
33	319.5	53.8	116	4	Q96PF6	Q96PF6 mus musculus
34	319.5	53.8	234	11	Q8VCP0	Q8VCP0 mus musculus
35	318.5	53.6	234	11	Q91WP8	Q91WP8 mus musculus
36	317	53.4	112	11	Q8KIF2	Q8KIF2 mus musculus
37	314	52.9	106	5	Q9U410	Q9U410 schistosoma
38	306.5	51.6	148	11	Q8K122	Q8K122 mus musculus
39	305.5	51.4	97	11	Q9JL76	Q9JL76 mus musculus
40	301.5	50.8	101	11	Q9JL78	Q9JL78 mus musculus
41	301	50.7	241	11	Q921A6	Q921A6 mus musculus
42	298	50.2	112	11	Q8KIF0	Q8KIF0 mus musculus
43	294	49.5	235	11	Q91W12	Q91W12 mus musculus
44	293.5	49.4	127	11	Q925S9	Q925S9 mus musculus
45	270.5	45.5	109	6	Q9N0W5	Q9N0W5 oryctolagus

ALIGNMENTS

RESULT 1
Q9ER29 PRELIMINARY: PRT: 107 AA.
ID Q9ER29
AC Q9ER29;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Anti human TNF-alpha light chain variable region (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
RT "Cloning and sequencing of the light chain fragment of variable region
genes of an anti-hTNF-a monoclonal antibody.";
RT J. Cell. Mol. Immunol. 12:21-26(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RT "Construction and sequencing of the single-chain antibody gene of a
human TNF-alpha specific monoclonal antibody.";
RT Ti 4 Chun i Ta Hsueh Hsueh Pao 19:373-376(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RT Submitted (May-2000) to the EMBL/Genbank/DBD databases.
RL EMBL; AF26753; AAG23804.1; -
DR HSSP; P80362; 1WTL.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00447; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 107 107
SO SEQUENCE 107 AA; 11784 MW; 2B15EBA6604A26C3 CRC64;

Query Match 70.5%; Score 418.5; DB 11; Length 107;
Best Local Similarity 76.6%; Pred. No. 2,6e-38;
Matches 82; Conservative 10; Mismatches 14; Indels 1; Gaps 1;

OY 4 LTQSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 63
DB 1 MTQSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
OY 64 VPDRTFTGSGSGTDFLTITSRVQVEDLAIYYCHQ-YLSSWTGSGTKL 109
DB 61 VPDRTFTGSGSGTDFLTITSRVQVEDLAIYYCHQ-YLSSWTGSGTKL 107

RESULT 2

OY 09UL78 PRELIMINARY; PRT; 109 AA.
AC 09UL78
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLREL. 23, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus";
RT Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035036; AAD56272.1; -
DR HSP; P80362; 1MPL.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11646 MW; 5F675C52EC7EE197 CRC64;

Query Match 66.0%; Score 392; DB 4; Length 109;
Best Local Similarity 66.7%; Pred. No. 2,2e-35;
Matches 76; Conservative 18; Mismatches 14; Indels 6; Gaps 2;

OY 1 DIQLTQSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
DB 1 EIVLTQSPGTLISLSPGERATLSCASQSV-----SSYLAWYQKPGQAPRLIYGAASR 55
OY 61 ESGVPRFTGSGSGTDFLTITSRVQVEDLAIYYCHQYLS-WTFGGTKLEIKR 113
DB 56 AITPDPFSGSGSGTDFLTITSRLEPEDCAVYCCQCGSSPLTFGGTKVEIKR 109

RESULT 3

OY 096SA9 PRELIMINARY; PRT; 107 AA.
AC 096SA9
DT 01-DEC-2001 (TREMBLREL. 19, Created)
DT 01-DEC-2001 (TREMBLREL. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLREL. 23, Last annotation update)
DE Anti-streptococcal/anti-myosin immunoglobulin kappa light chain variable region (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=98375893; PubMed=9712075;
RA Adderson E.E., Shikhan A.R., Ward K.E., Cunningham M.W.;
RT "Molecular analysis of polyclonal antibodies from
RT rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin
RT antibody v region genes";
RL J. Immunol. 161:2020-2031(1998).
DR EMBL; U96396; AAB68785.1; -
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11520 MW; 4BB43E9C5B577F16 CRC64;

Query Match 65.0%; Score 386; DB 4; Length 107;
Best Local Similarity 65.5%; Pred. No. 9,7e-35;
Matches 74; Conservative 17; Mismatches 16; Indels 6; Gaps 1;

OY 1 DIQLTQSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
DB 1 DIQLTQSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
OY 61 ESGVPRFTGSGSGTDFLTITSRVQVEDLAIYYCHQYLSWTGSGTKLEIKR 113
DB 55 QSGVPRFTGSGSGTDFLTITSRVQVEDLAIYYCHQYLSWTGSGTKLEIKR 107

RESULT 4

OY 09UL77 PRELIMINARY; PRT; 108 AA.
AC 09UL77
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLREL. 23, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus";
RT Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035037; AAD56273.1; -
DR HSP; P01607; 1REI.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11738 MW; C06681716CAD1E3 CRC64;

Query Match 64.6%; Score 383.5; DB 4; Length 108;
Best Local Similarity 64.9%; Pred. No. 1,8e-34;
Matches 74; Conservative 17; Mismatches 16; Indels 7; Gaps 2;

OY 1 DIQLTQSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
DB 1 DIQLTQSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNYLAWYQKPGQSPKLLIYMASTR 60
OY 61 ESGVPRFTGSGSGTDFLTITSRVQVEDLAIYYCHQ-YLSSWTGSGTKLEIKR 113

```

Db      55 QSGVPSRFSGSGSDTFLTTISLQPEDFAFYTCQGSYSTSWTFGEGRKVEIKR 108
       :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 5
O9UL70 PRELIMINARY; PRT; 108 AA.
AC O9UL70:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Myosin-reactive Immunoglobulin light chain variable region
   (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
   Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
   fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RT "fetus.";
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
   fetus.";
DR EMBL: AF035044; AAD56280.1; -.
DR HSSP; P01607; IREL.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 108 AA; 11633 MW; BYBEDC3EA1FCCA37 CRC64;

Query Match          63.2%; Score 375.5; DB 4; Length 108;
Best Local Similarity 65.8%; Pred. No. 1.4e-33;
Matches 75; Conservative 15; Mismatches 17; Indels 7; Gaps 2

OY 1 DIQLTPSSSLAVSAGENWTVSCSKSSQSVLSANHKNTLANWYQKRGPSKLIIYMASTR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 DIQMTPSSSSLASVSAGDVRTITCRASQGI-----SNLIANYQKRGKVPSPSLIYAATSL 54
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY 61 ESGVDPDRFTGGSGSGTDFTLTISRQVEDLAIFYCHQYLSS-WTEFGGTRLEIKR 113
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 QSGVPSRFSGSGSDTFLTTISLQPEDFAFYTCQGSYSTSWTFGEGRKVEIKR 108
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 6
O9UL81 PRELIMINARY; PRT; 107 AA.
AC O9UL81:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Myosin-reactive Immunoglobulin light chain variable region
   (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
   Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
   fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RT EMBL; AF035043; AAD56269.1; -.
DR HSSP; P01607; IREL.
```

Query Match	Best Local Similarity	62.8%	Score 373;	DB 4;	Length 107;
Matches 71;	Conservative 18;	Mismatches 18;	Indels 6;	Gaps 1;	
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.					
DR PROSITE: PS50835; IG_LIKE; 1.					
FT NON_TER 1					
FT NON_TER 107					
SO SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;					
QY 1 DIQLTGSPSSLAIVAGENVNTSCSSQSVLYSAHNKNTLANYQKPGQSPKLLIYMASTR 60	DR InterPro: IPR007110; IG_LIKE.				
DR InterPro: IPR003006; IG_MHC.					
DR InterPro: IPR003596; IG_V.					
DR Pfam: PF00047; Ig; 1.					
DR SMART: SM00406; IGV; 1.</					

[illegible]

Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
EMBL: BC019760; AAH19760.1; -
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003006

[illegible]

ID	Q8NEK1	PRELIMINARY;	PRT;	234	AA.
AC	Q8NEK1;				

DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC030813; ANH030813.1; ..
 DR InterPro: IPR007110; I9_1Ike.
 DR InterPro: IPR003357; I9_CL.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003596; I9_V.
 DR Pfam: PF00047; I9; 2.
 DR SMART: SM00407; IGc1.1.
 DR SMART: SM00406; IGV. 1.
 DR PROSITE: PSS0835; I9_LIKE; 2.
 DR PROSITE: PSS00290; I9_MHC; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 234 AA; 25530 MW; 63168DEFD8D132F8 CRC64;

Query Match	60.58;	Score	359.5;	DB	4;	Length	234;
Best Local Similarity	62.68;	Pred	NO	3	10-31		

Matches	72; Conservative	18; Mismatches	16; Indels	9; Gaps	3

DY 1 D I O L T G S S S L A V S G E N V T M S C K S S O S V L S A N K N K L A Y A Q O K P G S P K L I I Y W A S T R 60
Db :
21 E I V T S O S A T I S V S P E R A T I S C R A S G V ----- T S L A N Y Q O L P G S P R I A T Y I G A S S R 74

QY 61 E S G V P D F R T S G S G S T D F L T I S R V O E B L A I Y C H O V S W - P T G G S K L E I K R 113
Dy :
75 A S G V P A R F S G S G S T E F L T I S S L Q S E D F A N Y I Q O Y - N K P H T F G G S T K L I D K R 128

```
Search completed: October 7, 2003, 19:19:53
Job time : 114.61 secs
```


Matches	93; Conservative	5; Mismatches	5; Indels	0; Gaps	0;
Qy	1	DIOLTPSSSLAVSAGENVNTMSCKSSQSVLSSANHKNTLAWYQKQPGQSPKLLIYMASTR	60		
	:				
	:				
Db	1	NIMMTQSPFSLAVSAGEKVTMSCKSSQSVLSSNQNTLAWYQKQPGQSPKLLIYMASTR	60		
Qy	61	ESGVPDRFTGSGSGTDFTLTLSRQVDELAIYYCHQYSSMTF	103		
	:				
	:				
Db	61	ESGVPDRFTGSGSGTDFTLTLSRQVDELAIYYCHQYSSMTF	103		

RESULT 3

Ig kappa chain V region (12.5H VL) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
 C:Accession: S41393
 R:Margalitte, C.; Gilbert, D.; Brard, F.; Tron, F.
 Submitted to the EMBL Data Library, January 1994
 A:Description: Structural characterization of an (NZB X NZW)F1 mouse-derived Igm anti-DN
 A:Reference number: S41393
 A:Accession: S41393
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-112 <MAR>
 A:Cross-references: EMBL:Z29536
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match	82.88;	Score 492;	DB 2;	Length 112;
Best Local Similarity	83.98;	Pred NO	9	50-36.

Matches 94; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

[illegible]

```

OY      61  ESCVPPDRFPGSGSGDFTLTISRQVEDIAIYYCHQYLSSWTFGGGTLEIK 112
        : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      61  DSGVPPDRFPGSGSGDFTLTISRQVEDIAIYYCHQYLSSWTFGGGTLEIK 112

```

RESULT 4

antibody light chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000
C:Accession: S51147
C:Ride Krulff, J.; Boel, E.; Toonenberg, T.
submitted to the EMBL Data Library, January 1995
A:Description: Selection and application of human SCFV antibody fragments from a semi-syn-
A:Reference number: S51147
A:Accession: S51147
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <DEK>
A:Cross-references: EMBL:X83714; NTD:G633227; PTDN:CAA58689.1; PTD:G633228
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
P:16-96/Domain: Immunoglobulin homology <IMM>

Query Match	82.7%;	Score 491.5;	DB 2;	Length 120;
Best Local Similarity	80.7%;	Pred. NO. 1e-35;		
Matches 92;	Conservative 11;	Mismatches 10;	Indels 1;	Gaps 1

Oy 1 DIQLTSPSSSLAVSAGENVYTMCKSSQSVLYSANHKNTLAMYQQKPGGSPKLLIYNASTR 600
||:||||| ||| | ::::|||||:::||||||| |||||
Db 1 DIELTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNLTAMYYQKPGGPRLIIYNASTR 600

```

Oy      61  ESGVPDRFTSGSGGTDFLTLTISRQVEDALYYCHQYLSL-WTFEGGKTKLEIKR 113
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61  ESGVPDRFTSGSGGTDFLTLTISLQAEADVANYCQYYSLTMTWGQGTKEIKR 114

```

RESULT 5
S26040

RESULT 5

C:Species: Mus musculus (house mouse)
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 21-Jan-2000
C:Accession: S26040, S78098
R:Okamoto, M.; Honjo, T
Nucleic Acids Res. 18, 1895, 1990
A:Title: Nucleotide sequences of the gene/cDNA coding for anti-murine erythrocyte autoantibody.
#:Reference number: S09216; MUID:9024558; PMID:2356368

A;Molecule type: DNA

A;Cross-references: EMBL:X51742

R;Okamoto, M.

A:Reference number: S78098

A;Molecule type: DNA

A;Cross-references: EMBL:X51742; NID:g52697; PIDN:CAA36032.1; PID:g52698

A; Introns: 22/1

C;Keywords: heterotetramer; immunoglobulin

F;26-138/Product: Ig kappa chain (fragment) #status predicted <MAT>

F;41-121/Domain: Immunoglobulin homology <IMM>

Query Match	82.28;	Score 488;	DB 2;	Length 138;
01 48	01 48	01 48	01 48	01 48

Matches 92; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 DIQLTSPSSLAVSAGENVTMCKSSQSVLYLSANHKNTLAYOQKPGQSKLLIYASTR 60
 ||::||||||| |||||:||||| |||||||||
 Db 26 DIVMSQSPSSLTVSGEKVTMCKSSQSLYSNNQNKNTLAYOQKPGQSKLLIYASTR 85

```
QY      61 ESGVPDRFTGSSSGTDFLTILISRVQVEDLAIIYCHQYLSSMTFGGGTRLEIKR 113
        |||||
Db      86 ESGVDRFTGSSSGTDFLTILISSVKAEADLAVFYCQQYYSEFLTFGAGTRLEIKR 138
```

RESULT 6

Ig light chain V region (S107/VH11 group 1-4) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
 C:Accession: P0407
 R:Behar, S.M.; Lustgarten, D.L.; Corbet, S.; Scharff, M.D.
 J. Exp. Med. 173, 731-741, 1991
 A:Title: Characterization of somatically mutated S107 VH11-encoded anti-DNA autoantib
 A:Reference number: P03376; MUID:9114903; PMID:1900062
 A:Accession: P0407
 A:Molecule type: DNA
 A:Residues: 1-113

 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: Immunoglobulin
 C:KeyWords: Immunoglobulin homology <IM>
 C:16-96/Domain: Immunoglobulin homology

Query Match	82.0%;	Score 487;	DB 2;	Length 113;
Best Local Similarity	83.2%;	Pred. No. 2.3e-35;		
Matches 94; Conservative	6;	Mismatches 13;	Indels 0;	Gaps 0;

```

Oy      1 DIQLTQSPSSSLAVSAGENVMTSCCKSSQSVLYSANHKNTLYAQOKPGQSPKLLIYASTR 600
      ||:::||||| |||||:::| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 DIVISQSPSSSLAVSAGEKVTMCCKSSQSLNSTRFRKNTLYAQOKPGQSPKLLIYASTR 600

```

```
Oy      61 ESGVPDRFTSGSGTDFLLISRVQEDLAIYYCHQYLSNWTFGGGTLEIKR 113
        | | | | | | | | | | | | | | | : | | | | | | | |
Db      61 ECGVPDRFTSGSGTDFLLISSVQADLAIVCKQSYNLTFGGGTLEIKR 113
```


C:Species: Homo sapiens (man)
 C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 01-Dec-2000
 C:Accession: A01903; F61458
 R:Schneider, M.; Hilschmann, N.
 Hoppe-Sejler's Z. Physiol. Chem. 356, 507-557, 1975
 A:Title: Die Primerstruktur einer monoklonalen Immunglobulin-L-Kette der Subgruppe IV
 A:Reference number: A01903; MUID:76004342; PMID:50995
 A:Accession: A01903
 A:Molecule type: protein
 A:Residues: 1-114 <SCH>
 A>Note: this is the first completely sequenced V region of a new kappa chain subgroup, d
 R:Brown, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.; Minasco, E.
 J. Exp. Med. 170, 1551-1558, 1989
 A:Title: Expression of a public Idiotypic by human monoclonal IGM directed to myelin-assc
 A:Reference number: A61458; MUID:90039128; PMID:2478651
 A:Accession: F61458
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-13 <BRD>
 C:Comment: This is a Bence Jones protein.
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a
 C:Keywords: heterotetramer; immunoglobulin
 F:16-96/Domain: immunoglobulin homology <IMM>
 F:23-94/Disulfide bonds: #status predicted

Query Match 79.9%; Score 474.5; DB 1; Length 114;
 Best Local Similarity 78.9%; Pred. No. 2.8e-34;
 Matches 90; Conservative 12; Mismatches 11; Indels 1; Gaps 1;

OY 1 DIQLTQSPSSILAVSAGENVYMSCKSSQSVLYSANHKNTLAWYQQRPGQSPKLLIYMASTR 60
 ||:||||:||||:| ||:||||:||||:| ||:||||:||||:| ||:||||:||||:|
 DB 1 DIVMTQSPNSLAVSGERATINCKSSQSVLYSSNSKNTLAWYQQRPGQPKLLIYMASTR 60
 ||:||||:||||:| ||:||||:||||:| ||:||||:||||:| ||:||||:||||:|

OY 61 ESGVDRFTGSGSGDTFTLTISRVOEDLAIYCHQYLSS-WTPEGGTRLEIKR 113
 |||||:|||||:|||||:| ||:||||:| ||:||||:| ||:||||:| ||:||||:|
 DB 61 ESGVDRFTGSGSGDTFTLTISRVOEDLAIYCHQYLSS-WTPEGGTRLEIKR 114

Search completed: October 7, 2003, 19:21:47
 Job time : 43.0828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 22.9897 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-2

Perfect score: 594
Sequence: 1 DIQLTQSPSPSLAVSAGENVY.....CHQYLSSWTFGGGKLEIKR 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	80.8	133	1	KV4B_HUMAN
2	474.5	79.9	134	1	KV4C_HUMAN
3	473.5	79.7	114	1	KV4A_HUMAN
4	434	73.1	121	1	KV4D_HUMAN
5	420.5	70.8	109	1	KV4M_HUMAN
6	392	66.0	129	1	KV3M_HUMAN
7	386.5	65.1	114	1	KV3P_MOUSE
8	382	64.3	110	1	KV3L_MOUSE
9	382	64.3	129	1	KV3L_HUMAN
10	380.5	64.1	108	1	KV1M_HUMAN
11	379	63.8	129	1	KV3H_HUMAN
12	378	63.6	109	1	KV3B_HUMAN
13	377	63.5	109	1	KV3D_HUMAN
14	375	63.1	109	1	KV3E_HUMAN
15	369.5	62.2	108	1	KV1V_HUMAN
16	368.5	62.0	111	1	KV3O_MOUSE
17	368	62.0	133	1	KV2E_HUMAN
18	365.5	61.5	111	1	KV3L_MOUSE
19	365.5	61.5	136	1	KV5B_MOUSE
20	364.5	61.4	149	1	KV5A_MOUSE
21	364	61.3	117	1	KV2E_HUMAN
22	362.5	61.0	128	1	KV3K_HUMAN
23	362	60.9	108	1	KV3A_HUMAN
24	361.5	60.9	108	1	KV1J_HUMAN
25	361	60.8	109	1	KV3G_HUMAN
26	360.5	60.7	108	1	KV1H_HUMAN
27	360.5	60.7	108	1	KV1R_HUMAN
28	360.5	60.7	111	1	KV3O_MOUSE
29	360.5	60.7	115	1	KV2A_MOUSE
30	359.5	60.5	108	1	KV1B_HUMAN
31	359	60.4	109	1	KV3F_HUMAN
32	359	60.4	113	1	KV2G_MOUSE
33	358.5	60.4	129	1	KV1W_HUMAN

34	358	60.3	107	1	KV1D_HUMAN	P01596	homo sapien
35	358	60.3	113	1	KV2D_HUMAN	P01617	homo sapien
36	357.5	60.2	108	1	KV1K_HUMAN	P01603	homo sapien
37	357.5	60.2	111	1	KV3M_MOUSE	P01665	mus musculus
38	352.5	59.3	108	1	KV1F_MOUSE	P01598	homo sapien
39	352.5	59.3	108	1	KV1O_HUMAN	P01607	homo sapien
40	352.5	59.3	111	1	KV3N_MOUSE	P01666	mus musculus
41	351.5	59.2	111	1	KV3R_MOUSE	P01670	mus musculus
42	351.5	59.2	111	1	KV3U_MOUSE	P01673	mus musculus
43	351.5	59.2	131	1	KV3I_MOUSE	P01661	mus musculus
44	351	59.1	113	1	KV2B_HUMAN	P01615	homo sapien
45	350.5	59.0	112	1	KV2C_HUMAN	P01616	homo sapien

ALIGNMENTS

RESULT 1
KV4B_HUMAN STANDARD; PRT; 133 AA.
ID KV4B_HUMAN
AC P06313;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region JI precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid-9606;
RN [1]
RP MEDLINE-66041853; PubMed-2997712;
RA Klobbeck H.G.; Bornkamm G.W.; Combrato G.; Mochkat R.; Pohlenz H.D.;
RA Zacher H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene.";
RL Nucleic Acids Res. 13:6515-6529(1985);
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z00022; CAAT7317.1; -
DR PIR: A01904; KAHUT1.
DR HSP: P80362; IWT1.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003065; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20
FT 21 133 IG KAPPA CHAIN V-IV REGION JI.
FT 22 43 FRAMEWORK-1.
FT 23 60 COMPLEMENTARITY-DETERMINING-1.
FT 24 75 FRAMEWORK-2.
FT 25 82 COMPLEMENTARITY-DETERMINING-2.
FT 26 83 FRAMEWORK-3.
FT 27 114 COMPLEMENTARITY-DETERMINING-3.
FT 28 122 FRAMEWORK-4.
FT 29 132 BY SIMILARITY.
FT 30 133
FT 31 133
FT 32 133
FT 33 133
SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;
Query Match 80.8%; Score 480; DB 1; Length 133;


```

RESULT 4
KV40_HUMAN
ID KV40_HUMAN STANDARD: PRT: 121 AA.
AC P06312;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig kappa chain V-IV region precursor (Fragment).
GN IGV4-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP MEDLINE:86041853; PubMed:2997712;
RX Klobek H.G., Bornkamm G.W., Comarriato G., Mochkat R., Pohlentz H.D.,
RA Zaehun H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene."
RL Nucleic Acids Res. 13:6515-6529(1985).
CC -1 MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z00023; CAA77318.1; -.
DR PIR; A01902; KAHU.
DR HSSP; P80362; IMTL.
DR Genew; HGNC:5834; IGV4-1.
DR GO; GO:0005575; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003506; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_Like; 1.
DR Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 20 IG KAPPA CHAIN V-IV REGION.
FT SIGNAL 21 >121 FRAMEWORK-1.
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13380 MM; 9586AD418BD33974 CXC64;

Query Match 73.1%; Score 434; DB 1; Length 121;
Best Local Similarity 80.0%; Pred. No. 4e-39;
Matches 80; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIQLTSPSSILAVSAGENVMSCKSSOSVLYSANHKNTLAMYQOKPGSPKLLIYMASTR 60
DB 21 DIWVTQSPDLSILAVSLGERATINCKSSOSVLYSSNNKNTLAMYQOKPGSPKLLIYMASTR 80
QY 61 ESGVPDFRFTSGSGSCTDFTLTISRQVEDLAIYCHQYLS 100
DB 81 ESGVPDFRFTSGSGSCTDFTLTISRQVEDLAIYCHQYLS 120

```

ID	NAME	STANDARD	PRT	109 AA
RV4D	HUMAN			
AC	P83593			
DT	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig kappa chain V-IV region STR (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
NC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Abdominal adipose tissue;			
RX	PubMed=95588180;			
RA	Olsen K.E., Sletten K., Westermark P.;			
RT	"Extended analysis of AL-amyloid protein from abdominal wall			
RT	subcutaneous fat biopsies: kappa Ig immunoglobulin light chain.";			
RL	Biochem. Biophys. Res. Commun. 245:713-716(1998).			
CC	-1- FUNCTION: May play an important role in fibrillogenesis.			
DR	InterPro: IPR003559; Ig.			
DR	InterPro: IPR007110; Ig-like.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; Ig_1.			
DR	PROSITE: PS50835; Ig_LIKE; 1.			
KW	Immunoglobulin V region.			
FT	DOMAIN 1 23			
FT	DOMAIN 24 40			
FT	DOMAIN 41 55			
FT	DOMAIN 56 62			
FT	DOMAIN 63 94			
FT	DOMAIN 95 101			
FT	DOMAIN 102 109			
FT	DISULFID 23 94			
FT	UNSURE 23 23			
FT	UNSURE 94 94			
FT	NON_TER 109			
SO	SEQUENCE 109 AA; 12060 MW; 0C4F31EA11E12A0B CRC64;			
	Query Match 70.8%; Score 420.5; DB 1; Length 109;			
	Best Local Similarity 72.5%; Pred. No. 9.7e-38;			
	Matches 79; Conservative 12; Mismatches 17; Indels 1; Gaps 1;			
QY	1 DIOLTQSSSLAVSAGEVNTVMSCKSSQSVLYSAHKNTLAMYQKPGQSPFLITYMASTR 60			
DB	1 DIVATQSPDSLTVLSIGERATVINCSSQSVLYSSNNKNTLAMYQKPGQAPRLTFSMASTR 60			
QY	61 EGGVPRRTGSGSGNDFTLTISRQVEELATYYCHQVYS-SWTPGGGGR 108			
DB	61 EGGVPRRTGSGSGNDFTLTISRQVEELATYYCHQVYS-SWTPGGGGR 109			
	RESULT 6			
	KV3M_HUMAN			
ID	KV3M_HUMAN	STANDARD;	PRT; 129 AA.	
AC	P18136;			
DT	01-NOV-1980 (Rel. 16, Created)			
DT	01-NOV-1980 (Rel. 16, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig kappa chain V-III region HIC precursor.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
NC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88171307; PubMed=3127527;			
RA	Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;			
RT	"Anticentibody-associated kappa light chain variable region gene			
RT	expressed in chronic lymphocytic leukemia with little or no somatic			
RT	mutation. Implications for etiology and immunotherapy.";			
J.	J. Exp. Med. 167:840-852(1988).			
NC	-1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M			

AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC LEUKEMIA.

CC PIR: PL0021; K3HGH1.
 DR HSP: P80362; 1MTL.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; 1g; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 129
 FT DOMAIN 21 43 IG KAPPA CHAIN V-III REGION HIC.
 FT DOMAIN 44 55 FRAMEWORK-1.
 FT DOMAIN 56 70 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 71 77 FRAMEWORK-2.
 FT DOMAIN 78 109 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 110 118 FRAMEWORK-3.
 FT DOMAIN 119 129 COMPLEMENTARITY-DETERMINING-3.
 FT DISULFID 43 109 JKL SEGMENT.
 FT NON_TER 129 129 BY SIMILARITY.
 SQ SEQUENCE 129 AA; 14070 MW; 7395528EA2B74D6 CRC64;

Query Match 66.0%; Score 392; DB 1; Length 129;
 Best Local Similarity 65.8%; Pred. No. 1.2e-34;
 Matches 75; Conservative 18; Mismatches 15; Indels 6; Gaps 2;

QY 1 DIQLTQSPSSSLAVSAGENTVMSCKSSQSVLYSANHKNTLAWYQOKPGSPRLITYMASTR 60
 ID KVA1A_MOUSE STANDARD; PRT; 114 AA.
 AC P01632;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-I region S107A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81241357; PubMed=6788890;
 RA Kwan S.-P., Rudloff S., Seidman J.G., Leder P., Scharff M.D.;
 RT "Nucleic acid and protein sequences of phosphocholine-binding light
 chain";
 RT J. Exp. Med. 153:1366-1370(1981).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

RESULT 7
 KVA1A_MOUSE STANDARD; PRT; 114 AA.
 ID KVA1A_MOUSE STANDARD; PRT; 114 AA.
 AC P01632;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-I region S107A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81241357; PubMed=6788890;
 RA Kwan S.-P., Rudloff S., Seidman J.G., Leder P., Scharff M.D.;
 RT "Nucleic acid and protein sequences of phosphocholine-binding light
 chain";
 RT J. Exp. Med. 153:1366-1370(1981).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; 1g; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region.
 FT DOMAIN 1 23
 FT DOMAIN 24 40
 FT DOMAIN 41 55
 FT DOMAIN 56 62
 FT DOMAIN 63 94
 FT DOMAIN 95 103
 FT DOMAIN 104 113
 FT DISULFID 23 94
 FT NON_TER 114 114
 SQ SEQUENCE 114 AA; 12717 MW; 32008EC8E9DBE67B CRC64;

Query Match 65.1%; Score 386.5; DB 1; Length 114;
 Best Local Similarity 70.4%; Pred. No. 4e-34;
 Matches 81; Conservative 12; Mismatches 19; Indels 3; Gaps 3;

QY 1 DIQLTQSPSSSLAVSAGENTVMSCKSSQSVLYSANHKNTLAWYQOKPGSPRLITYMASTR 59
 ID KVA3P_MOUSE STANDARD; PRT; 110 AA.
 AC P01668;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-III region PC 7210.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79073152; PubMed=103003;
 RA Weigert M., Gaumattian L., Ioh E., Schilling J., Hood L.E.;
 RT "Rearrangement of genetic information may produce immunoglobulin
 RT diversity";
 RT Nature 276:785-790(1978).
 RL HSP: P01679; 2PRT.
 DR HSP: P01679; KVM510.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; 1g; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region.
 FT DOMAIN 1 23
 FT DOMAIN 24 38
 FT DOMAIN 39 53
 FT DOMAIN 54 60
 FT DOMAIN 61 92
 FT DOMAIN 93 100
 FT DOMAIN 101 110
 FT DISULFID 23 92
 FT NON_TER 110 110
 SQ SEQUENCE 110 AA; 11950 MW; 69F1A5CE886B1249 CRC64;

Query Match 64.3%; Score 382; DB 1; Length 110;
 Best Local Similarity 67.0%; Pred. No. 1.2e-33;
 Matches 75; Conservative 11; Mismatches 24; Indels 2; Gaps 1;

QY 1 DIQLTQSPSSSLAVSAGENTVMSCKSSQSVLYSANHKNTLAWYQOKPGSPRLITYMASTR 60
 ID KVA1A_MOUSE STANDARD; PRT; 114 AA.
 AC P01632;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-I region S107A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81241357; PubMed=6788890;
 RA Kwan S.-P., Rudloff S., Seidman J.G., Leder P., Scharff M.D.;
 RT "Nucleic acid and protein sequences of phosphocholine-binding light
 chain";
 RT J. Exp. Med. 153:1366-1370(1981).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

```

DB      1 DIVLTQASALAVSLGQARTTSCAKASQSIDYDGD--STANNYQOKRPGQPKLLITYAASNL 58
OY      61 ESGVDPRTGSGSGTDFLTLLISRVQVEDIAYYCHQYLSMTFGGKLEIKR 112
        59 ESGIAPRFGSGSGTDFLTLLIHVVEEDAAATYCHQSDPWFFGSTKLEIKR 110
DB      110

RESULT 9
KV3L_HUMAN  STANDARD; PRT; 129 AA.
ID   KV3L_HUMAN
AC   P18135;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DE   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-III region HAH precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88171307; PubMed=3127527;
RA   Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;
RT   "Autoantibody-associated kappa light chain variable region gene
RT   expressed in chronic lymphocytic leukemia with little or no somatic
RT   mutation. Implications for etiology and immunotherapy."
RL   J. Exp. Med. 167:840-852(1988).
CC   -1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M
CC   AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC
CC   LEUKEMIA.
DR   PIR: P10022; K3H0HA.
DR   HSSP: P80362; 1WTL.
DR   GO: GO:0005576; C:extracellular; NAS.
DR   GO: GO:0003823; F:antigen binding activity; NAS.
DR   GO: GO:0006955; P:immune response; NAS.
DR   InterPro: IPR007110; Ig-like.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam: PF00047; Ig_V.
DR   SMART: SM00406; IGV; 1.
DR   PROSITE: PS50835; IG_LIKE; 1.
DR   KX Immunoglobulin V region; Signal.
FT   SIGNAL 1 20
FT   CHAIN 21 129 IG KAPPA CHAIN V-III REGION HAH.
FT   DOMAIN 21 43 FRAMEWORK-1.
FT   DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN 56 70 FRAMEWORK-2.
FT   DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN 78 109 FRAMEWORK-3.
FT   DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN 119 129 JKL SEGMENT.
FT   DISULFID 43 109 BY SIMILARITY.
FT   NON_TER 129 129
SQ   SEQUENCE 129 AA; 14073 MW; D3C55292772774D0 CRC64;

Query Match 64.3%; Score 382; DB 1; Length 129;
Best Local Similarity 64.9%; Pred. No. 1.4e-33;
Matches 74; Conservative 18; Mismatches 16; Indels 6; Gaps 2;

```

```

DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-I region Lay.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=7038198; PubMed=824717;
RA   Capra J.D., Klapper D.G.;
RT   "Complete amino acid sequence of the variable domains of two human
RT   Igm anti-gamma globulins (Lay/Pom) with shared idiotypic
RT   specificities."
RL   Scand. J. Immunol. 5:677-684(1976).
CC   -1- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC   CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POK V-III KAPPA CHAIN,
CC   WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC   -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC   GLOBULIN ACTIVITY.
DR   PIR: A01871; K1H0LY.
DR   HSSP: P01607; 1REI.
DR   GO: GO:0005576; C:extracellular; NAS.
DR   GO: GO:0003823; F:antigen binding activity; NAS.
DR   GO: GO:0006955; P:immune response; NAS.
DR   InterPro: IPR007110; Ig-like.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam: PF00047; IGV; 1.
DR   SMART: SM00406; IGV; 1.
DR   PROSITE: PS50835; IG_LIKE; 1.
DR   KX Immunoglobulin V region.
FT   DOMAIN 1 23 FRAMEWORK-1.
FT   DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN 35 49 FRAMEWORK-2.
FT   DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN 57 88 FRAMEWORK-3.
FT   DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN 98 107 FRAMEWORK-4.
FT   DISULFID 23 88 BY SIMILARITY.
FT   NON_TER 108 108
SQ   SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;

Query Match 64.1%; Score 380.5; DB 1; Length 108;
Best Local Similarity 65.2%; Pred. No. 1.6e-33;
Matches 75; Conservative 17; Mismatches 14; Indels 9; Gaps 3;

```

RA Goldfien R., Carson D.A.;
 RT "Cloning and sequence determination of a human rheumatoid factor
 light-chain gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2195-2199(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.1sb-slb.ch/announce/>
 or send an email to license@slb-slb.ch).
 CC -----
 DR EMBL, M12740; AA58992.1; -
 DR HSSP; P80362; IWTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Immunoglobulin V region; Signal.
 KW SIGNAL.
 FT CHAIN 1 20
 FT 21 129 IG KAPPA CHAIN V-III REGION CLL.
 FT 21 43 FRAMEWORK-1.
 FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 55 69 FRAMEWORK-2.
 FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 77 108 FRAMEWORK-3.
 FT DOMAIN 109 118 COMPLEMENTARITY-DETERMINING-3.
 FT DISULFID 43 108 JKI SEGMENT.
 FT NON_TER 129 129 BY SIMILARITY.
 SO SEQUENCE 129 AA; 14275 MW; 5C13B411BE6CC14 CRC64;
 Query Match 63.8%; Score 379; DB 1; Length 129;
 Best Local Similarity 63.5%; Pred. No. 2.9e-33;
 Matches 73; Conservative 18; Mismatches 16; Indels 8; Gaps 2;
 QY 1 DIQLTQSPSSLAIVSAGENVMTSCSKSSQSVLYSANHKRYLAWYQKPGQSPKRLITYMASTR 60
 DB 21 EIVWTQSPGTLSPGSRATLSCRAASQV-----SNKLAWYQKPGQAPRRLLITYGASSTR 74
 QY 61 ESGVDPRTGSGCTDFTLTISRQVEDLAITYCHQYLS--SWTFGGTKLEIKR 113
 DB 75 ATGIPARFSGSGCTDFTLTISRQVEDLAITYCHQYLS--SWTFGGTKLEIKR 129
 RESULT 12
 KV3D_HUMAN STANDARD; PRT; 109 AA.
 ID KV3D_HUMAN
 AC P01620;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-III region SIE.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=82046598; PubMed=6794615;
 RA Andrews D.W., Capra J.D.;
 RT "Amino acid sequence of the variable regions of light chains from two
 idiotypically cross-reactive human Igm anti-gamma-globulins of the Wa
 group."
 RL Biochemistry 20:5816-5822(1981).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
 GLOBULIN ACTIVITY.

DR PIR; A01892; K3HUST.
 DR HSSP; P80362; IWTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Immunoglobulin V region.
 KW DISULFID 23 89 BY SIMILARITY.
 FT NON_TER 109 109
 SO SEQUENCE 109 AA; 11775 MW; 7689C3ECD646FFB4 CRC64;
 Query Match 63.6%; Score 378; DB 1; Length 109;
 Best Local Similarity 64.0%; Pred. No. 3.1e-33;
 Matches 73; Conservative 20; Mismatches 15; Indels 6; Gaps 2;
 QY 1 DIQLTQSPSSLAIVSAGENVMTSCSKSSQSVLYSANHKRYLAWYQKPGQSPKRLITYMASTR 60
 DB 1 EIVWTQSPGTLSPGSRATLSCRAASQV-----SNKLAWYQKPGQAPRRLLITYGASSTR 55
 QY 61 ESGVDPRTGSGCTDFTLTISRQVEDLAITYCHQYLS--WTFGGTKLEIKR 113
 DB 56 ATGIPARFSGSGCTDFTLTISRQVEDLAITYCHQYLS--WTFGGTKLEIKR 109
 RESULT 13
 KV3D_HUMAN STANDARD; PRT; 109 AA.
 ID KV3D_HUMAN
 AC P01622;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-III region TI.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=72188439; PubMed=5027703;
 RA Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein
 TI). IV. The complete amino acid sequence and its significance for
 the mechanism of antibody production."
 RL Hoppe-Seyler's Z. Physiol. Chem. 353:189-208(1972).
 CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
 CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
 DR PIR; A01895; K3HUT1.
 DR HSSP; P80362; IWTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region; Bence-Jones protein.
 FT DISULFID 23 89 BY SIMILARITY.
 FT NON_TER 109 109
 SO SEQUENCE 109 AA; 11788 MW; 8C35058CDCT749BC CRC64;
 Query Match 63.5%; Score 377; DB 1; Length 109;
 Best Local Similarity 64.0%; Pred. No. 3.9e-33;
 Matches 73; Conservative 20; Mismatches 15; Indels 6; Gaps 2;
 QY 1 DIQLTQSPSSLAIVSAGENVMTSCSKSSQSVLYSANHKRYLAWYQKPGQSPKRLITYMASTR 60

```
DB 1 EIVLQSGTSLSPGERATLSCRSQSV-----SNSLWYQOQPGAPRLLIYVASSR 55
OY 61 ESGVDRFTGSGSGDTFTLTISRQVEDLAIYCHQYLS-WTFGGGTRLEIKR 113
DB 56 ATGIPDRFSGSGSGDTFTLTISRLEPEFAVYYCOQYSSRSTFGGTRKVEIKR 109

RESULT 14
KV3E_HUMAN
ID KV3E_HUMAN STANDARD; PRT; 109 AA.
AC P01623;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region MOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=82046598; PubMed=6794615;
RA Andrews D.W., Capra J.D.;
RT *Antibody acid sequence of the variable regions of light chains from two
RT 1dotypically cross-reactive human IgM anti-gamma-globulins of the Wa
RT group.*;
RL Biochemistry 20:5816-5822(1981).
CC -1 MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC PIR: A01896; K3HMTL.
DR HSSP; P80362; 1MTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region.
FT DISULFD 23 89
FT NON_TER 109 109
FT SEQUENCE 109 AA; 11746 MW; 566C115E6B9CBEE CRC64;

Query Match 63.1%; Score 375; DB 1; Length 109;
Best Local Similarity 64.0%; Pred. No. 6.3e-33;
Matches 73; Conservative 17; Mismatches 18; Indels 6; Gaps 2;

OY 1 DIQLTQSPSSLAVSAGENVMTSCSSQSVLYSANHRNTLWYQOQPGSPKLLIYMASTR 60
DB 1 EIVLQSGTSLSPGERATLSCRSQSV-----SSGYLWYQOQPGAPRLLIYGASSR 55
OY 61 ESGVDRFTGSGSGDTFTLTISRQVEDLAIYCHQYLS-WTFGGGTRLEIKR 113
DB 56 ATGIPDRFSGSGSGDTFTLTISRLEPEFAVYYCOQYSSRSTFGGTRKVEIKR 109

RESULT 15
KV1V_HUMAN
ID KV1V_HUMAN STANDARD; PRT; 108 AA.
AC P04430;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region BAN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
```

```
RX MEDLINE=86174817; PubMed=3083240;
RA Dwyer F.E., O'Connor T.P., Benson M.D.;
RT *Polymorphism in a kappa I primary (AL) amyloid protein (BAN).*;
RL Mol. Immunol. 23:73-78(1986).
DR PIR: A01878; K1HJBN.
DR HSSP; P80362; 1MTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region; Amyloid.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 107 FRAMEWORK-4.
FT DISULFD 23 88
FT NON_TER 108 108
FT SEQUENCE 108 AA; 11840 MW; CD3FD944FE96FD37 CRC64;

Query Match 62.2%; Score 369.5; DB 1; Length 108;
Best Local Similarity 64.9%; Pred. No. 2.4e-32;
Matches 74; Conservative 15; Mismatches 18; Indels 7; Gaps 2;

OY 1 DIQLTQSPSSLAVSAGENVMTSCSSQSVLYSANHRNTLWYQOQPGSPKLLIYMASTR 60
DB 1 DIQLTQSPSSLAVSAGENVMTSCSSQSVLYSANHRNTLWYQOQPGAPRLLIYDASTL 54
OY 61 ESGVDRFTGSGSGDTFTLTISRQVEDLAIYCHQYLS-WTFGGGTRLEIKR 113
DB 55 QSGVSNFTGSGSGDTFTLTISRLEPEFAVYYCOQYSSRSTFGGTRKVEIKR 108

Search completed: October 7, 2003, 19:14:55
Job time : 22.9897 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:26 ; Search time 142.224 Seconds

(without alignments)
126.112 Million cell updates/sec

Title: US-09-988-013A-2

Perfect score: 594
Sequence: 1 DIQLTQSPSSLAIVSAGENV.....CHQYLSSWTFGGTKLEIKR 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	594	100.0	113	17	AA192215
2	594	100.0	113	18	AA192215
3	556	93.6	263	20	AA190226
4	556	93.6	268	20	AA190228
5	556	93.6	268	20	AA190222
6	556	93.6	273	20	AA190224
7	556	93.6	556	20	AA190218
8	556	93.6	580	20	AA190217
9	547	92.1	112	15	AA194933

10	540	90.9	112	23	AA192926
11	533	89.7	112	23	AA192926
12	531	89.4	112	23	AA192926
13	531	89.4	113	17	AA192217
14	530	89.2	272	23	AA191024
15	527	88.7	112	23	AA192927
16	523	88.0	115	18	AA192927
17	513	86.4	238	23	AA192927
18	508	85.5	112	23	AA192927
19	506.5	85.3	110	23	AA192927
20	504	84.8	130	24	AA192927
21	502.5	84.6	113	15	AA192927
22	502.5	84.6	251	22	AA192927
23	502.5	84.6	258	22	AA192927
24	502.5	84.6	259	21	AA192927
25	501	84.3	112	13	AA192807
26	499.5	84.1	134	22	AA192807
27	497	83.7	112	15	AA192927
28	497	83.7	301	18	AA192927
29	497	83.7	301	18	AA192927
30	497	83.7	301	22	AA192927
31	497	83.7	301	22	AA192927
32	497	83.7	352	20	AA192927
33	497	83.7	553	18	AA192927
34	497	83.7	553	20	AA192927
35	497	83.7	553	22	AA192927
36	497	83.7	553	22	AA192927
37	490.5	82.6	113	13	AA1925407
38	490.5	82.6	220	18	AA1925407
39	490	82.5	134	18	AA1925407
40	486.5	81.9	113	21	AA1925407
41	485.5	81.7	139	17	AA1925407
42	485	81.6	113	22	AA1925407
43	485	81.6	288	20	AA1925407
44	485	81.6	673	20	AA1925407
45	484.5	81.6	137	18	AA1925407

ALIGNMENTS

RESULT 1	AA192215	standard; Protein; 113 AA.
AC	AA192215;	
AC	AA192215;	
DT	28-MAY-1996	(first entry)
XX		
DE	LL2 MAB VR region.	
XX		
KW	Humanised antibody; monoclonal antibody; MAb; LL2; B-cell lymphoma; leukaemia; therapy; diagnosis; complementarity determining region; CDR; antibody engineering.	
XX		
OS	Mus musculus.	
XX		
FT	Key	Location/Qualifiers
FT	Region	24..40
FT		/label=CDP1
FT	Region	/note="claim 6, page 44"
FT		56..62
FT	Region	/label=CDR2
FT		/note="claim 7, page 44"
FT	Region	95..103
FT		/label=CDR3
FT		/note="claim 8, page 44"
PN	WO9604925-A1.	
XX		
PD	22-FEB-1996.	
XX		
PF	11-AUG-1995;	95WO-US09641.

S. aureus ClfA spe
S. aureus ClfA spe
S. aureus ClfA spe
Humanised LL2 MAB
Synthetic mouse fu
S. aureus ClfA spe
Variable kappa cha
Mouse/human chimera
Variable chain of
light chain variab
Escherichia coli 1
MAB NF52 light cha
Single chain antib
Single chain antib
Antiviral scfv-ant
5A8 VL. Synthetic
Anti-CA125 bifunct
Fc receptor human
Single chain, huma
Multispecific sing
Single chain human
Single chain human
Anti-Fc gamma rece
Single chain anti-
H22-anti-CEA antib
Bispecific single
Bispecific single
light chain variab
Anti-HGF receptor
Xenograft antibody
260F9 hybridoma VL
Monoclonal antibod
Human HIV-1 monocl
Fusion protein PNG
Plasmid pNG/55.15
Mouse MAB 15 light

XX 12-AUG-1994; 94US-0289576.
 XX (IMMU-) IMMUNOMEDICS INC.
 XX
 XX Hansen H, Leung S;
 XX WPI, 1996-139454/14.
 XX N-PSDB; AAT15802.
 DR
 PT Chimeric and humanised IL2 antibodies - used to produce conjugates
 PT for the therapy and diagnosis of B-cell lymphoma(s) and
 PT leukaemia(s).
 XX
 PS Claim 5; Page 35-36; 70pp; English.
 XX
 XX The complementarity determining regions (CDRs) of mouse monoclonal
 CC antibody (Mab) IL2 VK (AAR92215) and VH (AAR92216) regions were
 CC recombinantly linked to the framework sequences of human VK and VH
 CC regions, respectively, to give humanised IL2 VK (AAR92217) and VH
 CC (AAR92218). These were subsequently linked, respectively, to human
 CC kappa and IgG1 constant regions. A humanised Mab was obtained that
 CC retained the B-lymphoma and leukaemia cell targeting and
 CC internalisation characteristics of the parental IL2 Mab, and which
 CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
 CC cytostatic agent for therapeutic appln.
 CC
 XX
 SQ Sequence 113 AA;
 Query Match 100.0%; Score 594; DB 17; Length 113;
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIOLTPSSSLAVSAGENTYMSCKSSQSVLYSANHNKNTLAWYQKPGSPKLLIYWASR 60
 DB 1 DIOLTPSSSLAVSAGENTYMSCKSSQSVLYSANHNKNTLAWYQKPGSPKLLIYWASR 60
 QY 61 ESGVPDRFTGSGSGDTFTLTISRVOEDLAIYYCHQYLSWTFGGGTKEIKR 113
 DB 61 ESGVPDRFTGSGSGDTFTLTISRVOEDLAIYYCHQYLSWTFGGGTKEIKR 113
 RESULT 2
 ID AAM27695 standard; Protein; 113 AA.
 AC AAM27695;
 DT 14-APR-1998 (first entry)
 DE Variable kappa chain of MAb IL2.
 XX
 KW Variable kappa chain; B cell; monoclonal antibody; Mab; IL2;
 KW B cell lymphoma; lymphocytic leukaemia cell; murine;
 KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
 KW chronic lymphocytic leukaemia.
 XX
 OS Mus sp.
 OS Synthetic.
 OS
 FH Key 18..20 Location/Qualifiers
 FH Region /note= "potential N-linked glycosylation site"
 FT Region 24..40 /note= "complementarity determining region 1"
 FT Region 56..62 /note= "complementarity determining region 2"
 FT Region 95..102 /note= "complementarity determining region 3"
 FT /note= "complementarity determining region 3"
 XX
 XX WO9734632-A1.
 XX PD 25-SBP-1997.

XX 19-MAR-1997; 97WO-US04196.
 XX
 XX 20-MAR-1996; 96US-0013709.
 XX
 XX (IMMU-) IMMUNOMEDICS INC.
 XX
 XX Hansen H, Leung S, Qu Z;
 XX WPI, 1997-479995/44.
 XX N-PSDB; AAT88128.
 DR
 PT Monoclonal antibody engineered to contain glycosylation site - in
 PT non-Fc constant heavy or light chain region, useful to diagnose or
 PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
 XX
 PS Example 3; Fig 4A; 88pp; English.
 XX
 XX The present sequence is the variable kappa chain of the
 CC B cell specific monoclonal antibody (Mab) IL2, which contains an
 CC engineered tri-peptide N-glycan acceptor sequence. IL2 is a highly
 CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
 CC murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukaemia. The glycosylation site allows a label or
 CC therapeutic agent of increased size to be conjugated to the
 CC carbohydrate moiety, without affecting the Mab's binding affinity
 CC or specificity.
 CC
 XX
 SQ Sequence 113 AA;
 Query Match 100.0%; Score 594; DB 18; Length 113;
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DIOLTPSSSLAVSAGENTYMSCKSSQSVLYSANHNKNTLAWYQKPGSPKLLIYWASR 60
 DB 1 DIOLTPSSSLAVSAGENTYMSCKSSQSVLYSANHNKNTLAWYQKPGSPKLLIYWASR 60
 QY 61 ESGVPDRFTGSGSGDTFTLTISRVOEDLAIYYCHQYLSWTFGGGTKEIKR 113
 DB 61 ESGVPDRFTGSGSGDTFTLTISRVOEDLAIYYCHQYLSWTFGGGTKEIKR 113
 RESULT 3
 ID AAM90226 standard; Protein; 263 AA.
 AC AAM90226;
 DT 10-MAY-1999 (first entry)
 DE Anti-B7.2 monospecific triabody 1G10.
 XX
 KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KW CD86; T cell activation; inhibitory graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; triabody; antibody; 1G10.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - synthetic.
 OS
 FH Key 1..24 Location/Qualifiers
 FH Peptide /note= "pe1b signal peptide"
 FT Region 25..144 /note= "anti B7.2 Mab VH region"
 FT Peptide 145..257 /note= "anti B7.2 Mab VL region"
 FT Peptide 258..263 /note= "His6 tag"
 XX
 XX WO9858965-A2.

```

XX 30-DEC-1998.
XX 22-JUN-1998; 98WO-EP03791.
XX 20-JUN-1997; 97EP-0870092.
XX (INNO-) INNOGENETICS NV.
XX Bosman A, Buyse M, Lorre K, Sablon E;
XX WPI: 1999-105615/09.
XX N-PSDB: AAX01660.
XX
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX immune diseases including allograft rejection
XX
XX Example 7.3; Fig 34; 182pp; English.
XX
XX This polypeptide comprises a 1G10 monospecific triabody composed
XX of the VH region of anti-B7.2 monoclonal antibody (Mab) 1G10
XX joined to the VL region of 1G10. A triabody is a mono- a bi- or
XX a trispecific molecule recognising simultaneously e.g. two B7.2
XX and one B7.1 molecules. It has a rigid structure that prevents
XX simultaneous binding to the 3 targets. Each antigen-binding site
XX is formed by pairing of one VH and one VL domain from the same or
XX from two different polypeptides. The invention relates to novel
XX molecules, including triabodies, which can cross-link and/or
XX cross-react with the costimulatory molecules B7.1 and B7.2 expressed
XX on professional antigen-presenting cells, leading to the inhibition
XX of antigen-specific T cell activation. Methods are provided for
XX the production of such B7-binding molecules, and for their use in
XX the treatment or prevention of diseases of the immune system, in
XX particular graft rejection, graft versus host disease, allergy and
XX autoimmune diseases (claimed).
XX
XX Sequence 263 AA:
XX
XX Query Match 93.6%; Score 556; DB 20; Length 263;
XX Best Local Similarity 92.0%; Pred. No. 9e-38;
XX Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 DIQLTQSSSLAVSAGENVMTSCSSOSVLYSANHKNTLAWYQOKPGSPKILTYMASTR 60
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 145 DIQLTQSSSLAVSAGENVMTSCSSOSVLYSANHKNTLAWYQOKPGSPKILTYMASTR 204
XX
XX QY 61 ESGVPRFTGSGSGTDFLTITISRVQVEDLAIYYCHQYLSWTFGGGTLEIKR 113
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 205 ESGVPRFTGSGSGTDFLTITISRVQVEDLAIYYCHQYLSWTFGGGTLEIKR 257
XX
XX RESULT 4
XX AAW90228
XX ID AAW90228 standard; Protein: 268 AA.
XX
XX AC AAW90228;
XX
XX 10-MAY-1999 (first entry)
XX
XX Anti-B7.1/anti-B7.2 bispecific triabody II.
XX
XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
XX CD86; T cell activation; inhibitor; graft versus host disease;
XX transplant rejection; allograft rejection; autoimmune disease;
XX allergy; therapy; human; triabody; antibody; B7-24; 1G10.
XX
XX Chimeric - Mus sp.
XX OS Chimeric - synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..39
XX Peptide /note= "g3p signal peptide"
XX FT 40..155
XX Region

```

```

FT Peptide /note= "anti B7.1 Mab VH region"
FT 156..268
FT /note= "anti B7.2 Mab VL region"
XX
XX WO9858965-A2.
XX 30-DEC-1998.
XX 22-JUN-1998; 98WO-EP03791.
XX 20-JUN-1997; 97EP-0870092.
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman A, Buyse M, Lorre K, Sablon E;
XX WPI: 1999-105615/09.
XX N-PSDB: AAX01662.
XX
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX immune diseases including allograft rejection
XX
XX Example 7.3; Fig 38; 182pp; English.
XX
XX This polypeptide comprises a bispecific triabody composed of the VH
XX region of anti-B7.1 monoclonal antibody (Mab) B7-24 joined to the
XX VL region of of anti-B7.2 Mab 1G10. A triabody is a mono- a bi- or
XX a trispecific molecule recognising simultaneously e.g. two B7.2
XX and one B7.1 molecules. It has a rigid structure that prevents
XX simultaneous binding to the 3 targets. Each antigen-binding site
XX is formed by pairing of one VH and one VL domain from the same or
XX from two different polypeptides. The invention relates to novel
XX molecules, including triabodies, which can cross-link and/or
XX cross-react with the costimulatory molecules B7.1 and B7.2 expressed
XX on professional antigen-presenting cells, leading to the inhibition
XX of antigen-specific T cell activation. Methods are provided for
XX the production of such B7-binding molecules, and for their use in
XX the treatment or prevention of diseases of the immune system, in
XX particular graft rejection, graft versus host disease, allergy and
XX autoimmune diseases (claimed).
XX
XX Sequence 268 AA:
XX
XX Query Match 93.6%; Score 556; DB 20; Length 268;
XX Best Local Similarity 92.0%; Pred. No. 9.1e-38;
XX Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 DIQLTQSSSLAVSAGENVMTSCSSOSVLYSANHKNTLAWYQOKPGSPKILTYMASTR 60
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 156 DIQLTQSSSLAVSAGENVMTSCSSOSVLYSANHKNTLAWYQOKPGSPKILTYMASTR 215
XX
XX QY 61 ESGVPRFTGSGSGTDFLTITISRVQVEDLAIYYCHQYLSWTFGGGTLEIKR 113
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 216 ESGVPRFTGSGSGTDFLTITISRVQVEDLAIYYCHQYLSWTFGGGTLEIKR 268
XX
XX RESULT 5
XX AAW90222
XX ID AAW90222 standard; Protein: 268 AA.
XX
XX AC AAW90222;
XX
XX 10-MAY-1999 (first entry)
XX
XX Anti-B7.2 monospecific diabody 1G-10.
XX
XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
XX CD86; T cell activation; inhibitor; graft versus host disease;
XX transplant rejection; allograft rejection; autoimmune disease;
XX allergy; therapy; human; diabody; antibody; 1G-10.
XX
XX Chimeric - Mus sp.
XX OS Chimeric - synthetic.

```

```

XX Key Location/Qualifiers
FH Peptide 1..24
FT /note="pe1b signal peptide"
FT Region 25..144
FT Peptide /note="anti B7.2 MAb VH region"
FT /note="G4S flexible linker"
FT Region 150..262
FT /note="anti B7.2 MAb VL region"
FT Peptide 263..268
FT /note="His6 tag"
XX
XX WO9858965-A2.
XX
XX 30-DEC-1998.
XX
XX 22-JUN-1998; 98WO-EP03791.
XX
XX 20-JUN-1997; 97EP-0870092.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman A, Buyse M, Lorre K, Sablon E;
XX
XX WPI; 1999-105615/09.
XX
XX N-PSDB; AAX01656.
XX
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX immune diseases including allograft rejection
XX
XX Example 7.2: Fig 26; 182pp; English.
XX
XX This polypeptide comprises a 16-10 monospecific diabody composed
XX of the VH region of anti-B7.2 monoclonal antibody (MAb) 16-10
XX joined via a short, flexible linker to the VL region of 16-10.
XX Mono- or bispecific bivalent molecules are generated by shortening
XX the flexible linker sequence between the VH and VL of the anti-B7.1
XX scFv B7-24, the anti-B7.2 scFv 1610 and the scFv molecule with
XX dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
XX for bispecific molecules by cross-pairing the VH and VL domains
XX from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
XX B7.12/B7.12). The invention relates to novel molecules, including
XX diabodies, which can cross-link and/or cross-react with the
XX costimulatory molecules B7.1 and B7.2 expressed on professional
XX antigen-presenting cells, leading to the inhibition of antigen-
XX specific T cell activation. Methods are provided for the
XX production of such B7-binding molecules, and for their use in the
XX treatment or prevention of diseases of the immune system, in
XX particular graft rejection, graft versus host disease, allergy and
XX autoimmune diseases (claimed).
XX
XX Sequence 268 AA:
XX
XX Query Match 93.6%; Score 556; DB 20; Length 268;
XX Best Local Similarity 92.0%; Pred. No. 9.1e-38;
XX Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 DIQLTQSPSSSLAVSAGEENTMCKSSQSVLYSANHKNYLAWYQKRGSPKLLIYWASTR 60
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 150 DIELTQSPSSSLAVSAGEENTMCKSSQSVLYSANHKNYLAWYQKRGSPKLLIYWASTR 209
XX
XX QY 61 ESGVPRFTGSGSGTDETLTISRYVEDAIYCHQYLYLSSWTFGGTKLEIKR 113
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 210 ESGVPRFTGSGSGTDETLTISRYVEDAIYCHQYLYLSSWTFGGTKLEIKR 262
XX
XX RESULT 6
XX ID AAW90224 Standard; Protein; 273 AA.
XX AC AAW90224;
XX XX

```

```

DT 10-MAY-1999 (first entry)
DE Anti-B7.1/anti-B7.2 bispecific diabody II.
XX
XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
XX CD86; T cell activation; inhibitor; graft versus host disease;
XX transplant rejection; allograft rejection; autoimmune disease;
XX allergy; therapy; human; diabody; antibody; B7-24; 1610.
XX
XX Chimeric - Mus sp.
XX Chimeric - synthetic.
XX
XX Key Location/Qualifiers
FH Peptide 1..39
FT /note="g3p signal peptide"
FT Region 40..155
FT Peptide /note="anti B7.1 MAb VH region"
FT /note="G4S flexible linker"
FT Region 161..273
FT /note="anti B7.2 MAb VL region"
FT MISC-difference 21
FT /note="encoded by TCA"
XX
XX WO9858965-A2.
XX
XX 30-DEC-1998.
XX
XX 22-JUN-1998; 98WO-EP03791.
XX
XX 20-JUN-1997; 97EP-0870092.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman A, Buyse M, Lorre K, Sablon E;
XX
XX WPI; 1999-105615/09.
XX
XX N-PSDB; AAX01658.
XX
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX immune diseases including allograft rejection
XX
XX Example 7.2: Fig 30; 182pp; English.
XX
XX This polypeptide comprises a bispecific diabody composed of the VH
XX region of anti-B7.1 monoclonal antibody (MAb) B7-24 joined via a
XX short, flexible linker to the VL region of anti-B7.2 MAb 1610.
XX Mono- or bispecific bivalent molecules are generated by shortening
XX the flexible linker sequence between the VH and VL of the anti-B7.1
XX scFv B7-24, the anti-B7.2 scFv 1610 and the scFv molecule with
XX dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
XX for bispecific molecules by cross-pairing the VH and VL domains
XX from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
XX B7.12/B7.12). The invention relates to novel molecules, including
XX diabodies, which can cross-link and/or cross-react with the
XX costimulatory molecules B7.1 and B7.2 expressed on professional
XX antigen-presenting cells, leading to the inhibition of antigen-
XX specific T cell activation. Methods are provided for the
XX production of such B7-binding molecules, and for their use in the
XX treatment or prevention of diseases of the immune system, in
XX particular graft rejection, graft versus host disease, allergy and
XX autoimmune diseases (claimed).
XX
XX Sequence 273 AA:
XX
XX Query Match 93.6%; Score 556; DB 20; Length 273;
XX Best Local Similarity 92.0%; Pred. No. 9.3e-38;
XX Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 DIQLTQSPSSSLAVSAGEENTMCKSSQSVLYSANHKNYLAWYQKRGSPKLLIYWASTR 60
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 161 DIELTQSPSSSLAVSAGEENTMCKSSQSVLYSANHKNYLAWYQKRGSPKLLIYWASTR 220
XX

```

QY 61 ESGVDPDRFTGSGSDFTLTISRQVEDLAIYYCHQYLSWTFFGGTLEIKR 113
 DB 221 ESGVDPDRFTGSGSDFTLTISRQVEDLAIYYCHQYLSWTFFGGTLEIKR 273

RESULT 7
 AAW90218 standard; Protein: 556 AA.
 AAW90218;
 10-MAY-1999 (first entry)

XX De Bisppecific tetravalent antibody B1TAB1G10-B7-24H6.
 XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KM CD86; T cell activation; inhibitor; graft versus host disease;
 KM transplant rejection; allograft rejection; autoimmune disease;
 KM allergy; therapy; human; bisppecific tetravalent antibody; B1TAB;
 KM B1TAB1G10-B7-24H6.
 XX Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic.

XX Key Location/Qualifiers
 FH Region 1..120
 FT /note= "VH region anti B7.2 MAb"
 FT Peptide 121..135
 FT /note= "(G4S3) flexible linker"
 FT Region 136..248
 FT /note= "VL region anti B7.2 MAB"
 FT Region 249..259
 FT /note= "human IgG3 hinge region"
 FT Domain 260..285
 FT /note= "helix-turn-helix dimerisation domain"
 FT Domain 286..305
 FT /note= "human IgG3 hinge domain"
 FT Region 306..426
 FT /note= "VH region anti B7.1 MAB"
 FT Peptide 427..441
 FT /note= "(G4S3) flexible linker"
 FT Region 442..550
 FT /note= "VL region anti B7.1 MAB"
 FT Peptide 551..556
 FT /note= "His6 tag"

PN W09858965-A2.
 XX 30-DEC-1998.
 PD 22-JUN-1998; 98W0-EP03791.
 PE 20-JUN-1997; 97EP-0870092.
 PR (INNO-) INNOGENETICS NV.
 PA Bosman A, Buysé M, Torre K, Sablon E;
 PI WPI: 1999-105615/09;
 DR N-PSDB: AAX01652.
 DR XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 FT immune diseases including allograft rejection
 FT Example 7.1; Fig 18; 182pp; English.
 PS This polypeptide comprises the bisppecific tetravalent antibody
 CC B1TAB1G10-B7-24H6. The molecule consists of 4 scfvs, i.e. 2
 CC anti B7.1 scfvs and 2 anti B7.2 scfvs (tetravalency). One single
 CC B1TAB is a homodimer of 2 identical molecules, each containing both
 CC an anti B7.1 and anti B7.2 scfv (bisppecificity). An anti-B7.1 and
 CC and anti-B7.2 scfv are linked using a dimerisation domain (see

CC AAW90219), which drives the homodimerisation of the molecule. DNA
 CC (see AAX01652) encoding the B1TAB has been constructed to allow
 CC expression of the B1TAB in transformed E. coli cells. The B1TAB
 CC cross-links, and/or cross-reacts, with the costimulatory molecules
 CC B7.1 and B7.2 that are expressed on the membrane of professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. The invention relates to such
 CC B7-binding molecules, methods for their production, and their use
 CC for treating or preventing diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 556 AA:
 SQ Query Match 93.6%; Score 556; DB 20; Length 556;
 Best Local Similarity 92.0%; Pred. No. 1,8e-37;
 Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 DIQLTQSPSSLAVSAGENVNWSCKSSQSVLYSANHKRYLAWYQOKPGOSPRLIYMASTR 60
 DB 136 DIQLTQSPSSLAVSAGENVNWSCKSSQSVLYSANHKRYLAWYQOKPGOSPRLIYMASTR 195

QY 61 ESGVDPDRFTGSGSDFTLTISRQVEDLAIYYCHQYLSWTFFGGTLEIKR 113
 DB 196 ESGVDPDRFTGSGSDFTLTISRQVEDLAIYYCHQYLSWTFFGGTLEIKR 248

RESULT 8
 AAW90217 standard; Protein: 580 AA.
 AAW90217;
 10-MAY-1999 (first entry)

XX De Bisppecific tetravalent antibody B1TABB7-24-1G10H6.
 XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KM CD86; T cell activation; inhibitor; graft versus host disease;
 KM transplant rejection; allograft rejection; autoimmune disease;
 KM allergy; therapy; human; bisppecific tetravalent antibody; B1TAB;
 KM B1TABB7-24-1G10H6.
 XX Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /note= "pe1b signal peptide"
 FT Region 25..138
 FT /note= "VH region anti B7.1 MAB"
 FT Peptide 139..153
 FT /note= "(G4S3) flexible linker"
 FT Region 154..262
 FT /note= "VL region anti B7.1 MAB"
 FT Region 263..273
 FT /note= "human IgG3 hinge region"
 FT Domain 274..308
 FT /note= "helix-turn-helix dimerisation domain"
 FT Domain 309..319
 FT /note= "human IgG3 hinge domain"
 FT Region 320..446
 FT /note= "VH region anti B7.2 MAB"
 FT Peptide 447..461
 FT /note= "(G4S3) flexible linker"
 FT Region 462..574
 FT /note= "VL region anti B7.2 MAB"
 FT Peptide 575..580
 FT /note= "His6 tag"
 FT Misc-difference 261
 FT /note= "encoded by CNG"
 FT Misc-difference 322..327

/note- "codons for these amino acids are not present in the DNA sequence for B7.1 and B7.2 scFvs provided in the specification"

MO9858965-A2.

30-DEC-1998.

22-JUN-1998; 98WO-EPO3791.

20-JUN-1997; 97EP-0870092.

(INNO-) INNOGENETICS NV.

Bosman A, Buyse M, Lorre K, Sablon E;

WPI; 1999-105615/09.

N-PSDB; AAX01651.

New molecules which bind B7.1 and B7.2 - useful to prevent and treat immune diseases including allograft rejection

Example 7.1; Fig 16; 182pp; English.

This polypeptide comprises the bispecific tetravalent antibody B7.1-B7.2-1G10H6. The molecule consists of 4 scFvs, i.e. 2 anti B7.1 scFvs and 2 anti B7.2 scFvs (tetravalency). One single B7.1-B7.2 is a homodimer of 2 identical molecules, each containing both an anti B7.1 and anti B7.2 scFv (bispecificity). An anti-B7.1 and an anti-B7.2 scFv are linked using a dimerisation domain (see AAX01651), which drives the homodimerisation of the molecule. DNA expression of the B7.1-B7.2 has been constructed to allow cross-links, and/or cross-reacts, with the costimulatory molecules B7.1 and B7.2 that are expressed on the membrane of professional antigen-presenting cells, leading to the inhibition of antigen-specific T cell activation. The invention relates to such B7-binding molecules, methods for their production, and their use for treating or preventing diseases of the immune system, in particular graft rejection, graft versus host disease, allergy and autoimmune diseases (claimed).

Sequence 580 AA:

Query Match 93.6%; Score 556; DB 20; Length 580;

Best Local Similarity 92.0%; Pred. No. 1.9e-37;

Matches 104; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

1 DIQLTQSPSSSLAVSAGEVTMTSCSSQSVLYSANHKNYLAWYQQRKQSPKLLIYWA5TR 60

462 DIELTQSPSSSLAVSAGEVTMTCKSSQSVLYSNQKNYLAWYQQRKQSPKLLIYWA5TR 521

61 ESGVPDRFTGSGSGTDFTLTISRVQEDAIYYCHQYLSWTFGGGSKLEIKR 113

522 ESGVPDRFTGSGSGTDFTLTISRVQEDAIYYCHQYLSWTFGGGSKLEIKR 574

RESULT 9

AAR54933 standard; peptide; 112 AA.

AAR54933;

25-MAR-2003 (updated)

19-OCT-1994 (first entry)

MAB 022 VK chain.

Fc receptor; FcR; humanized antibody; hab; IgG; cancer; allergy; autoimmune disease; heteroantibody; bifunctional antibody; immunotoxin; CDR; complementarity determining region; VH; heavy chain variable region; VK; kappa chain variable region;

mononuclear phagocyte; PCR; polymerase chain reaction; primer; site-directed mutagenesis; monoclonal antibody; MAb.

MO9410332-A1.

11-MAY-1994.

04-NOV-1993; 93WO-US10384.

04-NOV-1992; 92GB-0023377.

(MEDA-) MEDAREX INC.

Carr FJ, Harris WJ, Tempest PR;

WPI; 1994-167486/20.

New humanised antibodies to Fc receptors - used for diagnosis or for treatment of e.g. cancer, allergies and infectious and auto-immune diseases

Disclosure; Page 23; 36pp; English.

Humanized antibodies (habs) for IgG Fc receptors on human phagocytes comprise the CDR of mouse monoclonal antibody 22 (from hybridoma CC 022MCL-1), VH chains from human IgG NEMM or KOL, and VK chains from Ig REI. Sequences are provided for mouse 022 VH (AAR54931), humanized NEMM-based VH (022 NMVH, AAR54929), humanized KOL-based VH (022 KLVH, AAR54930), mouse 022 VK (AAR54933) and humanized REI-based VK (022 HVVK, AAR54932). During hab production, VH and VK CDNA were PCR amplified using primers given in AAO65378-87. Mutagenesis of clone M13VHPCR2 KOL VH (L71R) was performed using oligos AAO65388-89. The habs can be used in heteroantibody, bifunctional antibody and immunotoxin production. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 112 AA:

Query Match 92.1%; Score 547; DB 15; Length 112;

Best Local Similarity 92.0%; Pred. No. 2.2e-37;

Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

1 DIQLTQSPSSSLAVSAGEVTMTSCSSQSVLYSANHKNYLAWYQQRKQSPKLLIYWA5TR 60

1 NIYMTQSPSSSLAVSAGEVTMTSCSSQSVLYSNQKNYLAWYQQRKQSPKLLIYWA5TR 60

61 ESGVPDRFTGSGSGTDFTLTISRVQEDAIYYCHQYLSWTFGGGSKLEIKR 112

61 ESGVPDRFTGSGSGTDFTLTISRVQEDAIYYCHQYLSWTFGGGSKLEIKR 112

RESULT 10

AAE29266 standard; Protein; 112 AA.

AAE29266;

27-JAN-2003 (first entry)

S. aureus ClfA specific monoclonal antibody 12-9VLA-1 protein.

Clumping factor A; ClfA; fibrinogen; fibrin; ClfA40; ClfA33; N3 protein; immunological; staphylococcal infection; impetigo; pneumonia; furuncle;

septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

Staphylococcus aureus.

Key Location/Qualifiers

Region 24..40 /note- "CDR1"

Region 56..62

[illegible]

FT	Region	24..40	"CDR1"
FT	Region	56..62	"CDR2"
FT	Region	95..102	"CDR3"
FT	Region	/note="CDR3"	
XX	WO200272600-A2.		
XX	19-SEP-2002.		
XX	28-JAN-2002; 2002WO-US02296.		
XX	26-JAN-2001; 2001US-264072P.		
XX	12-MAR-2001; 2001US-274611P.		
XX	18-JUN-2001; 2001US-298413P.		
XX	30-JUL-2001; 2001US-308116P.		
XX	(INHT-), INHIBITEX INC.		
XX	Pattl JM, Hutchins JT, Domanski P, Patel P, Hall A;		
XX	WPI; 2002-759834/82.		
XX	N-PSDB; AAD46863.		
XX	New anti-clumping factor A (ClfA) monoclonal antibody, useful for		
XX	treating or preventing Staphylococcus aureus infection e.g. wound		
XX	infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis		
XX	in a human or animal		
XX	Claim 11; Page 34; 80pp; English.		
XX	The invention relates to monoclonal antibody which binds the clumping		
XX	factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA		
XX	monoclonal antibody is useful for treating or preventing S. aureus		
XX	infection in a human or animal, and for inhibiting the binding of		
XX	Staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragmen		
XX	S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3		
XX	protein is useful for inducing an immunological response in a human or		
XX	animal. These staphylococcal infections include wound infections, sepsis		
XX	impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The		
XX	present sequence is Staphylococcus aureus ClfA specific monoclonal		
XX	antibody 13-2VLA-1 (variable light sequence) protein.		
XX	Sequence 112 AA;		
XX	Query Match 89.7%; Score 533; DB 23; Length 112;		
XX	Best Local Similarity 90.2%; Pred. No. 3e-36;		
XX	Matches 101; Conservative 5; Mismatches 6; Indels 0; Gaps		
XX	1 DIQLTQSPSSLAIVSAGENVYMSCKSSQSVLYSANRKNYLAAMYQOKPGOSPRLIYMASTR 60		
XX	1 NIMMTQSPSSLAIVSAGENVYMSCKSSQSVLYSANRKNYLAAMYQOKPGOSPRLIYMASTR 60		
XX	61 ESGVPRFTGSGSGDTFTLTISRYVEDLAIIYCHQYLSWTFGGGTRLEIK 112		
XX	61 ESGVPRFTGSGSGDTFTLTINSVQAEADLAIVYCHQYLSWTFGGGTRLEIK 112		
XX	RESULT 12		
XX	AAE29268		
XX	AAE29268 standard; Protein; 112 AA.		
XX	AAE29268;		
XX	27-JAN-2003 (first entry)		
XX	S. aureus ClfA specific monoclonal antibody 35-220VLD-4 protein.		
XX	Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;		
XX	immunological; staphylococcal infection; impetigo; pneumonia; furuncle;		
XX	septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.		

KW malignant condition: rheumatoid arthritis; myasthenia gravis; psoriasis;
 KW Grave's disease; Hashimoto's thyroiditis; type I diabetes mellitus;
 KW multiple sclerosis; systemic lupus erythematosus; Sjogrens syndrome;
 KW immune thrombocytopenic purpura; scleroderma; cancer; Chom's disease;
 KW ulcerative colitis; inflammatory bowel disease; immunological effector;
 KW cell mediated cytotoxicity; complement dependent cytotoxicity;
 KW complement fixation; mouse.
 XX Mus musculus.
 OS Synthetic.
 XX WO200256910-A1.
 XX 25-JUL-2002.
 PD 17-JAN-2002; 2002WO-US01487.
 PF 17-JAN-2001; 2001US-0765208.
 PR 17-JAN-2001; 2001US-0765208.
 PA (GENE-) GENE-CRAFT INC.
 PI Ledbetter JA, Hayden-Ledbetter M;
 XX WPI: 2002-599691/64.
 DR N-PSDB: ABR89856.
 XX New human binding domain-immunoglobulin fusion protein useful for
 PT treating a subject having or suspected of having a B-cell disorder or
 PT malignant condition e.g. rheumatoid arthritis
 XX Disclosure: Page 119; 136pp; English.
 PS The invention describes a binding domain-immunoglobulin fusion protein
 CC that is capable of at least one immunological activity, comprising a
 CC binding domain polypeptide fused to an immunoglobulin hinge region
 CC polypeptide capable of specifically binding to an antigen, or an
 CC immunoglobulin heavy chain CH2 or CH3 constant region polypeptide fused
 CC to the hinge region polypeptide or to the CH2 constant region
 CC polypeptide. The fusion protein is useful for treating a subject having
 CC or suspected of having a B-cell disorder or malignant condition e.g.
 CC rheumatoid arthritis, myasthenia gravis, Grave's disease, Hashimoto's
 CC thyroiditis, type I diabetes mellitus, multiple sclerosis, systemic lupus
 CC erythematosus, Sjogrens syndrome, immune thrombocytopenic purpura,
 CC psoriasis, scleroderma, cancer and inflammatory bowel disease such as
 CC Chom's disease and ulcerative colitis. The fusion protein retains the
 CC ability to participate in well known immunological effector activities
 CC including antibody dependent cell mediated cytotoxicity and/or complement
 CC fixation in complement dependent cytotoxicity, despite having structures
 CC that would not be expected to be capable of promoting the effector
 CC activities. It can be produced in substantial quantities that are
 CC typically greater than those routinely attained with single-chain
 CC antibody constructs. This is the amino acid sequence of a synthetic
 CC mouse immunoglobulin fusion protein.
 XX Sequence 272 AA;
 SQ
 Query Match 89.2%; Score 530; DB 23; Length 272;
 Best Local Similarity 87.5%; Pred. No. 1.2e-35;
 Matches 98; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
 QY 1 DIQLTQSPSSSLAVSAGENVMTSKSSQSVLYSANHKNTLAWYQKPGSPKLLIYMASTR 60
 DB 21 NIMTQSPSSSLAVSAGENVMTSKSSQSVLYSANHKNTLAWYQKPGSPKLLIYMASTR 80
 QY 61 ESGVPDRFTGSGSGTDFLTITISRVQVEDLAIYCHQYLSWTFGGGTLEIK 112
 DB 81 ESGVPDRFTGSGSGTDFLTITISRVQVEDLAIYCHQYLSWTFGGGTLEIK 132
 RESULT 15
 AAE29270
 ID AAE29270 standard: Protein: 112 AA.
 XX

AC AAE29270;
 XX 27-JAN-2003 (first entry)
 DR S. aureus ClfA specific monoclonal antibody 12-9VL-Hu protein.
 XX Clumping factor A: ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
 KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
 KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
 XX Staphylococcus aureus.
 OS
 XX Key Location/Qualifiers
 FH Region 24..40
 FT /note= "CDR1"
 FT 56..62
 FT /note= "CDR2"
 FT 95..102
 FT Region /note= "CDR3"
 XX WO200272600-A2.
 XX 19-SEP-2002.
 PD 28-JAN-2002; 2002WO-US02296.
 PF 26-JAN-2001; 2001US-264072P.
 PR 12-MAR-2001; 2001US-274611P.
 PR 18-JUN-2001; 2001US-296413P.
 PR 30-JUL-2001; 2001US-308116P.
 XX (INHI-) INHIBITEX INC.
 PA Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
 XX WPI: 2002-759834/82.
 DR N-PSDB: AAD46869.
 XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
 PT treating or preventing Staphylococcus aureus infection e.g. wound
 PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis
 PT in a human or animal
 XX Claim 11; Page 42; 80pp; English.
 PS The invention relates to monoclonal antibody which binds the clumping
 CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
 CC monoclonal antibody is useful for treating or preventing S. aureus
 CC infection in a human or animal, and for inhibiting the binding of
 CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
 CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
 CC protein is useful for inducing an immunological response in a human or
 CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus ClfA specific monoclonal
 CC antibody 12-9VL-Hu (humanised variable light sequence) protein.
 XX Sequence 112 AA;
 SQ
 Query Match 88.7%; Score 527; DB 23; Length 112;
 Best Local Similarity 88.4%; Pred. No. 9.2e-36;
 Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 QY 1 DIQLTQSPSSSLAVSAGENVMTSKSSQSVLYSANHKNTLAWYQKPGSPKLLIYMASTR 60
 DB 1 DIWMTQSPSSSLAVSAGENVMTSKSSQSVLYSANHKNTLAWYQKPGSPKLLIYMASTR 60
 QY 61 ESGVPDRFTGSGSGTDFLTITISRVQVEDLAIYCHQYLSWTFGGGTLEIK 112
 DB 61 ESGVPDRFTGSGSGTDFLTITISRVQVEDLAIYCHQYLSWTFGGGTLEIK 112
 Search completed: October 7, 2003, 19:13:51

Wed Oct 8 09:40:34 2003

Job time : 144.224 secs

us-09-988-013a-2.rag

Page 10

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 58.4483 Seconds
(without alignments)
305.878 Million cell updates/sec

Title: US-09-988-013a-2

Perfect score: 594
Sequence: 1 DIQLTGSPSSLAIVSAGEVNT.....CHQYLSWTFGGTKLEIKR 113

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCY_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCYUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	594	100.0	113	10	US-09-741-843-2
2	594	100.0	113	11	US-09-894-839-2
3	594	100.0	113	11	US-09-988-013a-2
4	547	92.1	112	12	US-10-229-335-28
5	540	90.9	112	15	US-10-056-052-10
6	533	89.7	112	15	US-10-056-052-6
7	531	89.4	112	15	US-10-056-052-14
8	530	89.2	272	12	US-10-053-530-14
9	530	89.2	272	15	US-10-053-530-14
10	527	88.7	112	15	US-10-056-052-18
11	525	88.4	113	10	US-09-741-843-6
12	525	88.4	113	11	US-09-894-839-6
13	525	88.4	113	11	US-09-988-013a-6
14	504	84.8	130	14	US-10-146-305-7
15	497	83.7	112	12	US-10-229-335-4

16	496	83.5	113	11	US-09-894-839-20	Sequence 20, Appl
17	490.5	82.6	220	10	US-09-995-693-1	Sequence 1, Appl1
18	490.5	82.6	220	15	US-10-232-408-1	Sequence 1, Appl1
19	484	81.5	132	10	US-09-249-011a-14	Sequence 4, Appl1
20	479	80.6	141	12	US-10-390-986-16	Sequence 16, Appl
21	478.5	80.6	113	12	US-10-160-506-71	Sequence 71, Appl
22	478.5	80.6	113	12	US-10-160-506-91	Sequence 91, Appl
23	474.5	79.9	113	12	US-09-215-163-42	Sequence 42, Appl
24	473.5	79.7	155	12	US-10-345-618-11	Sequence 11, Appl
25	473.5	79.7	142	12	US-10-345-618-6	Sequence 6, Appl1
26	473.5	79.7	495	12	US-10-345-618-4	Sequence 4, Appl1
27	471.5	79.4	134	10	US-10-255-478-58	Sequence 58, Appl
28	469.5	79.0	114	10	US-09-214-163E-2	Sequence 2, Appl1
29	469.5	79.0	114	10	US-09-214-163E-16	Sequence 6, Appl1
30	468.5	78.9	113	10	US-09-274-163E-16	Sequence 16, Appl
31	468.5	78.9	114	10	US-09-274-163E-4	Sequence 4, Appl1
32	468.5	78.9	284	12	US-10-255-478-70	Sequence 70, Appl
33	467.5	78.7	113	10	US-09-956-206A-80	Sequence 80, Appl
34	467.5	78.7	113	12	US-10-255-478-73	Sequence 73, Appl
35	467.5	78.7	240	9	US-09-799-514-8	Sequence 8, Appl1
36	466.5	78.5	114	9	US-09-810-502-38	Sequence 38, Appl
37	466.5	78.5	220	12	US-10-159-006-17	Sequence 17, Appl
38	466.5	78.5	240	12	US-10-159-006-28	Sequence 28, Appl
39	465.5	78.5	274	12	US-10-255-478-66	Sequence 66, Appl
40	463.5	78.0	115	10	US-09-998-831-9	Sequence 9, Appl1
41	463.5	78.0	115	10	US-10-373-561-9	Sequence 9, Appl1
42	462.5	77.9	113	10	US-09-999-025-7	Sequence 7, Appl1
43	462.5	77.9	113	10	US-09-999-025-13	Sequence 13, Appl
44	462.5	77.9	113	10	US-09-999-040-7	Sequence 7, Appl1
45	462.5	77.9	113	10	US-09-999-040-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-741-843-2
Sequence 2, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
FILE REFERENCE: 018733/0996
CURRENT APPLICATION NUMBER: US/09/741.843
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-741-843-2

Query Match 100.0%; Score 594; DB 10; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.1e-47;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIQLTGSPSSLAIVSAGEVNTMSCKSSQSVLSANHKNTLAMYQOKPGQSPRLTIYMASTR 60
DB 1 DIQLTGSPSSLAIVSAGEVNTMSCKSSQSVLSANHKNTLAMYQOKPGQSPRLTIYMASTR 60
QY 61 ESGVPDFGTGSGSDFTLTISRQVEDLAIYVHQTLSWTGCGTKLEIKR 113
DB 61 ESGVPDFGTGSGSDFTLTISRQVEDLAIYVHQTLSWTGCGTKLEIKR 113

```
RESULT 2
US-09-894-839-2
; Sequence 2, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-2

Query Match          100.0%; Score 594; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIQLQSPSSLA VSA GENTMCKSSQSVLYSANHNKNTLAWYQOKGSPKLLIYASTR 60
    |||
DB 1 DIQLQSPSSLA VSA GENTMCKSSQSVLYSANHNKNTLAWYQOKGSPKLLIYASTR 60

QY 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 113
    |||
DB 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 113

RESULT 3
US-09-988-013A-2
; Sequence 2, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-988-013A-2

Query Match          100.0%; Score 594; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DIQLQSPSSLA VSA GENTMCKSSQSVLYSANHNKNTLAWYQOKGSPKLLIYASTR 60
    |||
DB 1 DIQLQSPSSLA VSA GENTMCKSSQSVLYSANHNKNTLAWYQOKGSPKLLIYASTR 60
```

```
QY 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 113
    |||
DB 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 113

RESULT 4
US-10-229-335-28
; Sequence 28, Application US/10229335
; Publication No. US20030144483A1
; GENERAL INFORMATION:
; APPLICANT: MEDAREX, INC.
; TITLE OF INVENTION: HUMANIZED ANTIBODIES TO FC RECEPTORS FOR
; IMMUNOBLOBULIN G ON HUMAN MONONUCLEAR PHAGOCYTES
; NUMBER OF SEQUENCES: 28
; STREET: P.O. Box 953, 1545 Route 22 East
; CITY: Annandale
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/229,335
; FILING DATE: 26-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,516
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-10-229-335-28

Query Match          92.1%; Score 547; DB 12; Length 112;
Best Local Similarity 92.0%; Pred. No. 4,4e-43;
Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIQLQSPSSLA VSA GENTMCKSSQSVLYSANHNKNTLAWYQOKGSPKLLIYASTR 60
    |||
DB 1 NIVMTQSPSSLA VSA GENTMCKSSQSVLYSNCKNTLAWYQOKGSPKLLIYASTR 60

QY 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 112
    |||
DB 61 ESGVDPRTFTGSGSGTDFLTITISRVQEDLAIYCHQYLSSWTFGGGKLEIKR 112

RESULT 5
US-10-056-052-10
; Sequence 10, Application US/10056052
; Publication No. US2003009656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: HUTCHINS, Jeff T.
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
US-10-056-052-10
```

FILE REFERENCE: P07069US04/BAS
CURRENT APPLICATION NUMBER: US/10/056,052
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/308,116
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/298,413
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/274,611
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/264,072
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 112
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-056-052-10

Query Match 90.9%; Score 540; DB 15; Length 112;
Best Local Similarity 91.1%; Pred. No. 1.9e-42;
Matches 102; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 DIOLQSPSSILAVSAGENVMTSCSSOSVLSANKHNTLAWYQOKPGSPKLLIYMASTR 60
DB 1 NIMMTQSSSLAVSAGEVMTSCSSOSVLSNKNLAWYQOKPGSPKLLIYMASTR 60
YQ 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112
DB 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112

RESULT 6
US-10-056-052-6
Sequence 6, Application US/10056052
Publication No. US2003009656A1
GENERAL INFORMATION:
APPLICANT: PATTI, Joseph M
APPLICANT: HUTCHINS, Jeff T
APPLICANT: DOMANSKI, Paul
APPLICANT: PATEL, Pratiksha
APPLICANT: HALL, Andrea
TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLPA PROTEIN . . .
FILE REFERENCE: P07069US04/BAS
CURRENT APPLICATION NUMBER: US/10/056,052
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/308,116
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/298,413
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/274,611
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/264,072
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 112
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-056-052-6

Query Match 89.7%; Score 533; DB 15; Length 112;
Best Local Similarity 90.2%; Pred. No. 8.4e-42;
Matches 101; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 DIOLQSPSSILAVSAGENVMTSCSSOSVLSANKHNTLAWYQOKPGSPKLLIYMASTR 60
DB 1 NIMMTQSSSLAVSAGEVMTSCSSOSVLSNKNLAWYQOKPGSPKLLIYMASTR 60
YQ 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112
DB 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112

RESULT 7
US-10-056-052-14
Sequence 14, Application US/10056052
Publication No. US2003009656A1
GENERAL INFORMATION:
APPLICANT: PATTI, Joseph M
APPLICANT: HUTCHINS, Jeff T
APPLICANT: DOMANSKI, Paul
APPLICANT: PATEL, Pratiksha
APPLICANT: HALL, Andrea
TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLPA PROTEIN . . .
FILE REFERENCE: P07069US04/BAS
CURRENT APPLICATION NUMBER: US/10/056,052
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/308,116
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/298,413
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/274,611
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/264,072
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 112
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-056-052-14

Query Match 89.4%; Score 531; DB 15; Length 112;
Best Local Similarity 89.3%; Pred. No. 1.3e-41;
Matches 100; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 1 DIOLQSPSSILAVSAGENVMTSCSSOSVLSANKHNTLAWYQOKPGSPKLLIYMASTR 60
DB 1 NIMMTQSSSLAVSAGEVMTSCSSOSVLSNKNLAWYQOKPGSPKLLIYMASTR 60
YQ 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112
DB 61 ESGVPRFTSGSGTDFLTITISRYOVEDLAIYCHQYLSWTFFGGTLEIK 112

RESULT 8
US-10-053-530-14
Sequence 14, Application US/10053530
Publication No. US20030133939A1
GENERAL INFORMATION:
APPLICANT: Ledbetter, Jeffrey
APPLICANT: Hayden-Ledbetter, Martha
TITLE OF INVENTION: Binding Domain-Immunoglobulin Fusion Proteins
FILE REFERENCE: 390069,401
CURRENT APPLICATION NUMBER: US/10/053,530
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: US 09/765,208
PRIOR FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 272
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(272)
OTHER INFORMATION: MOOSE ANTI-HUMAN CD22 SCFV
US-10-053-530-14

Query Match 89.2%; Score 530; DB 12; Length 272;
Best Local Similarity 87.5%; Pred. No. 4e-41;
Matches 98; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

;; SEQ ID NO 6
;; LENGTH: 113
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-894-839-6

Query Match 88.4%; Score 525; DB 11; Length 113;
Best Local Similarity 85.0%; Pred. No. 4.6e-41;
Matches 96; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIOLTPSSSLAVSAGENVMTSCSSQSVLSANHNKNTLAWYQKPGSPKLLIYMASTR 60
1 DIOLTPSSSLASAVGDBVTMCKSSQSVLSANHNKNTLAWYQKPGAPKLLIYMASTR 60

DB 61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113
61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113

RESULT 13
US-09-988-013A-6
; Sequence 6, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 1873/1082
; CURRENT APPLICATION NUMBER: US/09/988, 013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741, 843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127, 902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690, 102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289, 576
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-988-013A-6

Query Match 88.4%; Score 525; DB 11; Length 113;
Best Local Similarity 85.0%; Pred. No. 4.6e-41;
Matches 96; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIOLTPSSSLAVSAGENVMTSCSSQSVLSANHNKNTLAWYQKPGSPKLLIYMASTR 60
1 DIOLTPSSSLASAVGDBVTMCKSSQSVLSANHNKNTLAWYQKPGAPKLLIYMASTR 60

DB 61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113
61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113

RESULT 14
US-10-146-305-7
; Sequence 7, Application US/10146305
; Publication No. US20020173035A1
; GENERAL INFORMATION:
; APPLICANT: YUHAN CORPORATION
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV
; TITLE OF INVENTION: S-SURFACE ANTIGEN AND A GENE ENCODING THE SAME
; FILE REFERENCE: OVI7440
; CURRENT APPLICATION NUMBER: US/10/146, 305
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: KR 10-2001-26634
; PRIOR FILING DATE: 2001-05-16

;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: Kopatentin 1.71
;; SEQ ID NO 7
;; LENGTH: 130
;; TYPE: PRF
;; ORGANISM: Escherichia coli
US-10-146-305-7

Query Match 84.8%; Score 504; DB 14; Length 130;
Best Local Similarity 85.0%; Pred. No. 4.5e-39;
Matches 96; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 DIOLTPSSSLAVSAGENVMTSCSSQSVLSANHNKNTLAWYQKPGSPKLLIYMASTR 60
1 DIVQSPSSSLAVSAGENVMTSCSSQSVLSANHNKNTLAWYQKPGSPKLLIYMASTR 60

DB 61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113
61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 113

RESULT 15
US-10-229-335-4

; Sequence 4, Application US/10229335
; Publication No. US2003014483A1
; GENERAL INFORMATION:
; APPLICANT: MEDAREX, INC.
; TITLE OF INVENTION: HUMANIZED ANTIBODIES TO FC RECEPTORS FOR
IMMUNOBLOBULIN G ON HUMAN MONONUCLEAR PHAGOCYTES

NUMBER OF SEQUENCES: 28
STREET: P.O. Box 953, 1545, Route 22 East
CITY: Annandale
STATE: New Jersey
COUNTRY: USA
ZIP: 08801

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/229, 335
FILING DATE: 26-Aug-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/435, 516
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MXI-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-229-335-4

Query Match 83.7%; Score 497; DB 12; Length 112;
Best Local Similarity 81.2%; Pred. No. 1.7e-38;
Matches 91; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIOLTPSSSLAVSAGENVMTSCSSQSVLSANHNKNTLAWYQKPGSPKLLIYMASTR 60
1 DIOLTPSSSLASAVGDBVTITCKSSQSVLSANHNKNTLAWYQKPGAPKLLIYMASTR 60

DB 61 ESGVPRRTGSGSGTDFLTLSRQVEDLAIYCHQYLSWTFFGGGTLEIKR 112

Db 61 ESCVPSRFSGSGGDDFTFTISSLOPEDIAITYCHOYLSSWTFGGGTVEIK 112

Search completed: October 7, 2003, 19:24:23
Job time : 59.4483 secs

XX 12-AUG-1994; 94US-0289576.
 PR (IMMU-) IMMUNOMEDICS INC.
 PA
 XX
 PI Hansen H, Leung S;
 XX WPI: 1996-139454/14.
 DR N-PSDB; AAT15802.
 XX
 PT Chimeric and humanised LL2 antibodies - used to produce conjugates
 PT for the therapy and diagnosis of B-cell lymphoma(s) and
 PT leukaemia(s).
 PS
 PS Claim 5; Page 35-36; 70pp; English.
 CC The complementarity determining regions (CDRs) of mouse monoclonal
 CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
 CC recombinantly linked to the framework sequences of human VK and VH
 CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
 CC (AAR92218). These were subsequently linked, respectively, to human
 CC kappa and IgG1 constant regions. A humanised Mab was obtained that
 CC retained the B-lymphoma and leukaemia cell targeting and
 CC internalisation characteristics of the parental LL2 Mab, and which
 CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
 CC cytostatic agent for therapeutic appln.
 CC
 SQ Sequence 113 AA;
 XX
 Query Match 100.0%; Score 85; DB 17; Length 113;
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KSSQSVLYSANHKNYLA 17
 Db 24 KSSQSVLYSANHKNYLA 40
 XX
 RESULT 2
 AAR92217
 ID AAR92217 standard; Protein; 113 AA.
 AC AAR92217;
 XX
 XX 28-MAY-1996 (first entry)
 DT
 XX
 DE Humanised LL2 MAB VK region.
 XX
 KW Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
 KW leukaemia; therapy; diagnosis; complementarity determining region;
 KW CDR; antibody engineering.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX FH Region 24..40
 XX FT /label- CDR1
 FT Region 56..62
 FT /label- CDR2
 FT 95..103
 FT /label- CDR3
 XX
 XX W09604925-A1.
 XX
 XX 22-FEB-1996.
 PD
 XX
 PF 11-AUG-1995; 95MO-US09641.
 XX
 PR 12-AUG-1994; 94US-0289576.
 XX
 PA (IMMU-) IMMUNOMEDICS INC.
 PA Hansen H, Leung S;
 PI

XX WPI: 1996-139454/14.
 DR N-PSDB; AAT15803.
 XX
 PI Chimeric and humanised LL2 antibodies - used to produce conjugates
 PT for the therapy and diagnosis of B-cell lymphoma(s) and
 PT leukaemia(s).
 PS
 PS Claim 5; Page 38; 70pp; English.
 CC The complementarity determining regions (CDRs) of mouse monoclonal
 CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
 CC recombinantly linked to the framework sequences of human VK and VH
 CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
 CC (AAR92218). These were subsequently linked, respectively, to human
 CC kappa and IgG1 constant regions. A humanised Mab was obtained that
 CC retained the B-lymphoma and leukaemia cell targeting and
 CC internalisation characteristics of the parental LL2 Mab, and which
 CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
 CC cytostatic agent for therapeutic appln.
 CC
 SQ Sequence 113 AA;
 XX
 Query Match 100.0%; Score 85; DB 17; Length 113;
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KSSQSVLYSANHKNYLA 17
 Db 24 KSSQSVLYSANHKNYLA 40
 XX
 RESULT 3
 AAW27695
 ID AAW27695 standard; Protein; 113 AA.
 AC AAW27695;
 XX
 XX 14-APR-1998 (first entry)
 DT
 XX
 DE Variable kappa chain of Mab LL2.
 XX
 XX Variable kappa chain; B cell; monoclonal antibody; Mab; LL2;
 KW B cell lymphoma; lymphocytic leukaemia cell; murine;
 KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
 KW chronic lymphocytic leukaemia.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX FH Region 18..20
 XX FT /note- "potential N-linked glycosylation site"
 XX FT Region 24..40
 XX FT /note- "complementarity determining region 1"
 FT Region 56..62
 FT /note- "complementarity determining region 2"
 FT 95..102
 FT /note- "complementarity determining region 3"
 XX
 XX W09734632-A1.
 XX
 XX 25-SEP-1997.
 PD
 XX
 PF 19-MAR-1997; 97MO-US04196.
 XX
 PR 20-MAR-1996; 96US-0013709.
 XX
 PA (IMMU-) IMMUNOMEDICS INC.
 PA Hansen H, Leung S; Qu Z;
 PI WPI: 1997-479995/44.
 DR

DR N-PSDB; AAT88128.
 XX Monoclonal antibody engineered to contain glycosylation site - in
 PT non-Fc constant heavy or light chain region, useful to diagnose or
 XX treat B cell malignancies, e.g. non-Hodgkins lymphoma
 PS Example 3; Fig 4A; 88pp; English.
 XX
 CC The present sequence is the variable kappa chain of the
 CC B cell specific monoclonal antibody (Mab) LL2, which contains an
 CC engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly
 CC specific anti-B cell lymphoma and anti-lymphocytic leukemia cell
 CC murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukemia. The glycosylation site allows a label or
 CC therapeutic agent of increased size to be conjugated to the
 CC carbohydrate moiety, without affecting the Mab's binding affinity
 CC or specificity.
 CC
 SQ Sequence 113 AA;

Query Match 100.0%; Score 85; DB 18; Length 113;
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYXSANKNYLA 17
 |||||||
 DB 24 KSSQSVLYXSANKNYLA 40

RESULT 4
 AAM27697
 ID AAM27697 standard; Protein: 115 AA.

AC AAM27697;

DT 14-Apr-1998 (first entry)

DE Variable kappa chain of Mab hLL2.

XX Variable kappa chain; B cell; monoclonal antibody; Mab; hLL2;
 KW B cell lymphoma; lymphocytic leukemia cell; murine; humanised;
 KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
 KW chronic lymphocytic leukemia.

XX ChimERIC - Mus sp.
 OS ChimERIC - Homo sapiens.

PH Key Location/Qualifiers

FT Region 24..40 /note= "complementarity determining region 1"

FT Region 56..62 /note= "complementarity determining region 2"

FT Region 95..102 /note= "complementarity determining region 3"

XX WO9734632-A1.

XX 25-SEP-1997.

XX 19-MAR-1997; 97MO-US04196.

XX 20-MAR-1996; 96US-0013709.

XX (IMMUNOMEDICS=INC)

XX Hansen H, Leung S, Qu Z;

XX WPI: 1997-479995/44.

XX N-PSDB; AAT88130.

PT Monoclonal antibody engineered to contain glycosylation site - in
 non-Fc constant heavy or light chain region, useful to diagnose or

PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
 XX
 PS Example 3; Fig 5A; 88pp; English.
 XX
 CC The present sequence is the variable kappa chain of the
 CC B cell specific monoclonal antibody (Mab) hLL2. hLL2 is a highly
 CC specific anti-B cell lymphoma and anti-lymphocytic leukemia cell
 CC humanised murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukemia.
 CC
 SQ Sequence 115 AA;

Query Match 100.0%; Score 85; DB 18; Length 115;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYXSANKNYLA 17
 |||||||
 DB 24 KSSQSVLYXSANKNYLA 40

RESULT 5

AAU70334
 ID AAU70334 standard; Peptide: 17 AA.

AC AAU70334;

DT 14-FEB-2002 (first entry)

DE Human kappa IV light chain CDRL.

XX Immunoglobulin; antibody; light chain; heavy chain; CDR; FR;
 KW complementarity determining region; framework region; IgBP;
 KW transgenic plant; immunoglobulin binding protein array;
 KW IgM; IgG; IgA; IgD; IgE; IgY; IgW; kappa; lambda; CHBP.

XX Homo sapiens.

XX WO200183806-A1.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US14349.

XX 02-MAY-2000; 2000US-0563222.

XX (EPIC-) EPICYTE PHARM INC.

XX Hiatt AC, Hein MB;

XX WPI: 2002-055482/07.

PT Preparing immunoglobulin binding protein array in plant cells by
 transforming the cells with different polynucleotides encoding binding
 protein polypeptides specific to ligand, selecting plant cells for
 preparing array

XX Disclosure; Page 14; 129pp; English.

XX The invention relates to transforming a population of cells (e.g. plant
 cells), comprising using a library of two different polynucleotides
 encoding different immunoglobulin binding protein (IgBP) polypeptides
 that specifically bind to a ligand or form one or more disulphide bonds
 with polypeptides in transfected cells, to generate an IgBP that
 binds to a ligand, and transformed plant cells are selected, and
 preparing an IgBP array in plant cells. At least one peptide sequence has
 at least 75% sequence identity to a framework region (FR) of a native
 IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule.
 The method is useful for preparing an immunoglobulin binding protein
 array, preferably heavy chain binding protein (CHBP) array in eukaryotic
 cells especially plant cells (e.g. Agrobacterium tumefaciens or maize)
 or other eukaryotic cells (e.g. insect cells or mammalian cells). The

CC CHBP is useful for discovery of e.g. screening assays of IgBPs having
CC desired characteristics. The present sequence is a mammalian
CC immunoglobulin derived peptide that may be incorporated into an IgBP of
CC the invention.

XX Sequence 17 AA;

Query Match

Best Local Similarity 88.2%; Score 75; DB 23; Length 17;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLYA 17
| | | | | | | | | | | | | | | | | |
DB 1 KSSQSVLYSSNNKNTLYA 17

RESULT 6
AAM62807
ID AAM62807 standard; Peptide; 82 AA.

XX AAM62807;

DT 23-SEP-1998 (first entry)

DE Amino acid sequence of a human antibody fragment.

XX Human; immunoglobulin; Ig; transgenic; non-human mammal;

KW inactivated endogenous Ig locus; B-cell development;
KW human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
KW production; antibody.

XX Homo sapiens.

PN WO9824893-A2.

PD 11-JUN-1998.

PF 03-DEC-1997; 97WO-US23091.

PR 03-DEC-1996; 96US-0759620.

XX (ABGE-) ABGENIX INC.

PI Green L, Jakobovits A, Klapholz S, Kucherlapati R;

XX Mendez M;

DR WPI; 1998-33314/29.

XX New transgenic non-human mammals - having an inactivated
PT immunoglobulin locus and a near complete human immunoglobulin locus,
PT used for production of human antibodies

XX Disclosure; Page 78; 128pp; English.

CC AAM62793-822 represent fragments of human antibodies produced by
CC transgenic Xenomice, created using the method of the invention. The
CC specification describes a transgenic non-human mammal which has genome
CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
CC locus, so that the mammal does not display normal B-cell development. The
CC modified genome also has an inserted human heavy chain Ig locus in
CC germline configuration, the human heavy chain Ig locus comprising a human
CC micro constant region and regulatory and switch sequences, human J-H
CC genes, human D-H genes, and human V-H genes and an inserted human kappa
CC light chain Ig locus in germline configuration, the human kappa light
CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
CC are selected to restore normal B-cell development in the mammal. The
CC transgenic animals have a near complete human Ig locus, including both a
CC human heavy chain locus and a human kappa light chain locus. They can
CC be used for the production of human antibodies when exposed to
CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha

CC respectively.

XX Sequence 82 AA;

Query Match

Best Local Similarity 88.2%; Score 75; DB 19; Length 82;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLYA 17
| | | | | | | | | | | | | | | | | |
DB 5 KSSQSVLYSSNNKNTLYA 21

RESULT 7
ABJ18696
ID ABJ18696 standard; Protein; 101 AA.

XX ABJ18696;

DT 06-MAR-2003 (first entry)

DE Antibody library related human protein sequence SEQ ID No 25.

KW Library; recombinant antibody; clustering variable region; in silico;
KW immunogenicity; antibody therapeutic; human.

XX Homo sapiens.

PN WO200284277-A1.

PD 24-OCT-2002.

PF 17-APR-2002; 2002WO-US12202.

PR 17-APR-2001; 2001US-284407P.

XX (ABMA-) ABMAXIS INC.

PI Luo P;

DR WPI; 2003-093043/08.

XX Constructing a library of recombinant antibodies useful as source of
PT antibody candidates for screening antigens comprises clustering
PT variable regions of antibodies having known 3-dimensional structures
PT into structural ensembles

XX Disclosure; Page 115; 119pp; English.

CC The invention relates to a novel method for the construction of a library
CC of recombinant antibodies. The novel method comprises clustering variable
CC regions of a collection of antibodies having known 3D structures into at
CC least two families of structural ensembles, each comprising at least two
CC different antibody sequences but with substantially identical main chain
CC conformations. The method is useful for constructing a library of
CC artificial antibodies in silico which provides a structurally diverse and
CC yet functionally more relevant source of antibody candidates which can
CC then be screened for binding a wide variety of target molecules,
CC including small molecules, and biomacromolecules such as proteins,
CC peptides and nucleic acids. The libraries constructed are useful as a
CC source of antibody candidates for further screening for novel antibodies
CC with high affinity against a wide range of antigens and having no or
CC minimum immunogenicity to human subjects treated with antibody
CC therapeutics. This sequence represents a human peptide region of an
CC antibody relating to the novel antibody library construction method of
CC the invention.

XX Sequence 101 AA;

Query Match

Best Local Similarity 88.2%; Score 75; DB 24; Length 101;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
 DB 24 KSSQSVLYSSNNKNYLA 40

RESULT 8

ID AAB18857 standard; Protein; 113 AA.

XX AAB18857;

DT 08-FEB-2001 (first entry)

DE Amino acid sequence of anti-p53 antibody light chain clone 163.2.

XX p53; antibody; immune response; vaccine; gene therapy; cancer;

OS Homo sapiens.

PN WO200056770-A1.

PD 28-SEP-2000.

PF 15-MAR-2000; 2000MO-AU00189.

PR 19-MAR-1999; 99AU-0009321.

PA (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.

PI Ward RL, Coomber DMJ;

PS WPI; 2000-638249/61.

DR N-PSDB; AAA96134.

XX Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide

PT fragments, useful in treatment and diagnosis of cancer, rheumatoid

PT arthritis and coronary heart disease

PS Claim 30; Page 132; 163pp; English.

XX The present sequence represents the light chain of an antibody reactive

CC against p53. The antibody is obtained from a vertebrate host expressing

CC an immune response against a naturally occurring disease. The antibodies

CC are useful in pharmaceutical compositions, which additionally contain

CC chelators, drugs, produgs, toxins and imaging markers e.g. radioisotopes

CC or gadolinium. The polypeptide components of the antibodies are useful in

CC vaccines, for inducing an immune response against a disease in a

CC vertebrate, for treatment and/or prophylaxis of disease and for detection

CC purposes. The nucleic acid sequences can be used to detect a disease as

CC well as for gene therapy and recombinant production of the polypeptides.

CC In particular, the following can be treated cancer, rheumatoid arthritis

CC and coronary heart disease. Cancers include carcinogenic tumours, tumours

CC of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,

CC head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,

CC gastric cancer, brain cancer, bladder cancer, prostate cancer and

CC urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,

CC e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.

XX Sequence 113 AA;

Query Match 88.2%; Score 75; DB 21; Length 113;

Best Local Similarity 88.2%; Pred. No. 3.5e-05;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17

DB 24 KSSQSVLYSSNNKNYLA 40

RESULT 9

AAB18861

ID AAB18861 standard; Protein; 113 AA.

XX AAB18861;

DE Amino acid sequence of anti-p53 antibody light chain clone 163.6.

XX p53; antibody; immune response; vaccine; gene therapy; cancer;

OS Homo sapiens.

PN WO200056770-A1.

PD 28-SEP-2000.

PF 15-MAR-2000; 2000MO-AU00189.

PR 19-MAR-1999; 99AU-0009321.

PA (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.

PI Ward RL, Coomber DMJ;

PS WPI; 2000-638249/61.

DR N-PSDB; AAA96138.

XX Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide

PT fragments, useful in treatment and diagnosis of cancer, rheumatoid

PT arthritis and coronary heart disease

PS Claim 30; Page 136; 163pp; English.

XX The present sequence represents the light chain of an antibody reactive

CC against p53. The antibody is obtained from a vertebrate host expressing

CC an immune response against a naturally occurring disease. The antibodies

CC are useful in pharmaceutical compositions, which additionally contain

CC chelators, drugs, produgs, toxins and imaging markers e.g. radioisotopes

CC or gadolinium. The polypeptide components of the antibodies are useful in

CC vaccines, for inducing an immune response against a disease in a

CC vertebrate, for treatment and/or prophylaxis of disease and for detection

CC purposes. The nucleic acid sequences can be used to detect a disease as

CC well as for gene therapy and recombinant production of the polypeptides.

CC In particular, the following can be treated cancer, rheumatoid arthritis

CC and coronary heart disease. Cancers include carcinogenic tumours, tumours

CC of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,

CC head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,

CC gastric cancer, brain cancer, bladder cancer, prostate cancer and

CC urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,

CC e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.

XX Sequence 113 AA;

Query Match 88.2%; Score 75; DB 21; Length 113;

Best Local Similarity 88.2%; Pred. No. 3.5e-05;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17

DB 24 KSSQSVLYSSNNKNYLA 40

RESULT 10

AAB18863

ID AAB18863 standard; Protein; 113 AA.

XX AAB18863;

DT 08-FEB-2001 (first entry)

DE Amino acid sequence of anti-p53 antibody light chain clone 163.7.

XX p53; antibody; immune response; vaccine; gene therapy; cancer;

XX

```

KW rheumatoid arthritis; coronary heart disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 101
FT /note= "Arg encoded by CCA"
XX
XX WO200056770-A1.
XX
XX 28-SEP-2000.
XX
XX PD
XX PF 15-MAR-2000; 2000WO-AU00189.
XX
XX PR 19-MAR-1999; 99AU-0009321.
XX
XX (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX
XX PA Ward RL, Coomber DMJ;
XX
XX PI
XX PS WPI: 2000-638249/61.
XX
XX DR N-PSDB: AAA96140.
XX
XX PT Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide
XX fragments, useful in treatment and diagnosis of cancer, rheumatoid
XX arthritis and coronary heart disease
XX
XX PS Claim 30; Page 138; 163pp; English.
XX
XX
XX The present sequence represents the light chain of an antibody reactive
XX against p53. The antibody is obtained from a vertebrate host expressing
XX an immune response against a naturally occurring disease. The antibodies
XX are useful in pharmaceutical compositions, which additionally contain
XX chelators, drugs, produgs, toxins and imaging markers e.g. radioisotopes
XX or gadolinium. The polypeptide components of the antibodies are useful in
XX vaccines, for inducing an immune response against a disease in a
XX vertebrate, for treatment and/or prophylaxis of disease and for detection
XX purposes. The nucleic acid sequences can be used to detect a disease as
XX well as for gene therapy and recombinant production of the polypeptides.
XX In particular, the following can be treated cancer, rheumatoid arthritis
XX and coronary heart disease. Cancers include carcinogenic tumours, tumours
XX of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,
XX head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,
XX gastric cancer, brain cancer, bladder cancer, prostate cancer and
XX urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,
XX e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.
XX
XX SQ Sequence 113 AA:
XX
XX Query Match 88.2%; Score 75; DB 21; Length 113;
XX Best Local Similarity 88.2%; Pred. NO. 3.5e-05;
XX Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KSSOSVLYSANHKNTLA 17
XX |||||||:|||||
XX Db 24 KSSOSVLYSSNKNKNTLA 40
XX
XX
XX RESULT 11
XX AAB1865
XX ID AAB1865 standard; Protein: 113 AA.
XX
XX AC AAB1865;
XX
XX DT 08-FEB-2001 (first entry)
XX
XX DE Amino acid sequence of anti-p53 antibody light chain clone 163.9.
XX
XX KW p53; antibody; immune response; vaccine; gene therapy; cancer;
XX rheumatoid arthritis; coronary heart disease.
XX
XX OS Homo sapiens.
XX

```

```

PN WO200056770-A1.
XX
XX PD 28-SEP-2000.
XX
XX PF 15-MAR-2000; 2000WO-AU00189.
XX
XX PR 19-MAR-1999; 99AU-0009321.
XX
XX (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX
XX PA Ward RL, Coomber DMJ;
XX
XX PI
XX PS WPI: 2000-638249/61.
XX
XX DR N-PSDB: AAA96142.
XX
XX PT Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide
XX fragments, useful in treatment and diagnosis of cancer, rheumatoid
XX arthritis and coronary heart disease
XX
XX PS Claim 30; Page 140; 163pp; English.
XX
XX
XX The present sequence represents the light chain of an antibody reactive
XX against p53. The antibody is obtained from a vertebrate host expressing
XX an immune response against a naturally occurring disease. The antibodies
XX are useful in pharmaceutical compositions, which additionally contain
XX chelators, drugs, produgs, toxins and imaging markers e.g. radioisotopes
XX or gadolinium. The polypeptide components of the antibodies are useful in
XX vaccines, for inducing an immune response against a disease in a
XX vertebrate, for treatment and/or prophylaxis of disease and for detection
XX purposes. The nucleic acid sequences can be used to detect a disease as
XX well as for gene therapy and recombinant production of the polypeptides.
XX In particular, the following can be treated cancer, rheumatoid arthritis
XX and coronary heart disease. Cancers include carcinogenic tumours, tumours
XX of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,
XX head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,
XX gastric cancer, brain cancer, bladder cancer, prostate cancer and
XX urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,
XX e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.
XX
XX SQ Sequence 113 AA:
XX
XX Query Match 88.2%; Score 75; DB 21; Length 113;
XX Best Local Similarity 88.2%; Pred. NO. 3.5e-05;
XX Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KSSOSVLYSANHKNTLA 17
XX |||||||:|||||
XX Db 24 KSSOSVLYSSNKNKNTLA 40
XX
XX
XX RESULT 12
XX AAB1867
XX ID AAB1867 standard; Protein: 113 AA.
XX
XX AC AAB1867;
XX
XX DT 08-FEB-2001 (first entry)
XX
XX DE Amino acid sequence of anti-p53 antibody light chain clone 163.14.
XX
XX KW p53; antibody; immune response; vaccine; gene therapy; cancer;
XX rheumatoid arthritis; coronary heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200056770-A1.
XX
XX PD 28-SEP-2000.
XX
XX PF 15-MAR-2000; 2000WO-AU00189.
XX
XX PR 19-MAR-1999; 99AU-0009321.
XX

```

PA	(SV1N-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX	
PI	Ward RL, Coomber DMW;
DR	WPI: 2000-638249/61.
DR	N-PSDB: AAA96144.
XX	
PT	Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide
PT	fragments, useful in treatment and diagnosis of cancer, Rheumatoid
PT	arthritis and coronary heart disease
XX	
PS	Claim 30; Page 142; 163pp; English.
XX	
CC	The present sequence represents the light chain of an antibody reactive
CC	against p53. The antibody is obtained from a vertebrate host expressing
CC	an immune response against a naturally occurring disease. The antibodies
CC	are useful in pharmaceutical compositions, which additionally contain
CC	chelators, drugs, prodrugs, toxins and imaging markers e.g. radioisotopes
CC	or gadolinium. The polypeptide components of the antibodies are useful in
CC	vaccines, for inducing an immune response against a disease in a
CC	vertebrate, for treatment and/or prophylaxis of disease and for detection
CC	purposes. The nucleic acid sequences can be used to detect a disease as
CC	well as for gene therapy and recombinant production of the polypeptides.
CC	In particular, the following can be treated cancer, rheumatoid arthritis
CC	and coronary heart disease. Cancers include carcinogenic tumours, tumours
CC	of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,
CC	head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,
CC	gastric cancer, brain cancer, bladder cancer, prostate cancer and
CC	urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,
CC	e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.
XX	
SO	Sequence 113 AA:
	Query Match 88.2%; Score 75; DB 21; Length 113;
	Best Local Similarity 88.2%; Pred. No. 3.5e-05;
	Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY	1 KSSQSVLYXSANKNYLA 17
	:::
DB	24 KSSQSVLYSSNNKYLA 40
RESULT 13	
AAB18869	
ID	AAB18869 standard; Protein; 113 AA.
XX	
AC	AAB18869;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Amino acid sequence of anti-p53 antibody light chain clone 163.15.
DE	
XX	
KW	p53; antibody; immune response; vaccine; gene therapy; cancer;
KW	rheumatoid arthritis; coronary heart disease.
XX	
OS	Homo sapiens.
XX	
PN	W0200056770-A1.
XX	
PD	28-SEP-2000.
XX	
PF	15-MAR-2000; 2000WO-AU00189.
XX	
PR	19-MAR-1999; 99AU-0009321.
XX	
PA	(SV1N-) ST VINCENT'S HOSPITAL SYDNEY LTD.
PI	Ward RL, Coomber DMW;
XX	
DR	WPI: 2000-638249/61.
DR	N-PSDB: AAA96146.
XX	
PT	Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide

```

PT fragments, useful in treatment and diagnosis of cancer, rheumatoid
PT arthritis and coronary heart disease
PS Claim 30; Page 144; 163pp; English.
XX
XX
XX The present sequence represents the light chain of an antibody reactive
CC against p53. The antibody is obtained from a vertebrate host expressing
CC an immune response against a naturally occurring disease. The antibodies
CC are useful in pharmaceutical compositions, which additionally contain
CC chelators, drugs, produgs, toxins and imaging markers e.g. radioisotopes
CC or gadolinium. The polypeptide components of the antibodies are useful in
CC vaccines, for inducing an immune response against a disease in a
CC vertebrate, for treatment and/or prophylaxis of disease and for detection
CC purposes. The nucleic acid sequences can be used to detect a disease as
CC well as for gene therapy and recombinant production of the polypeptides.
CC In particular, the following can be treated cancer, rheumatoid arthritis
CC and coronary heart disease. Cancers include carcinogenic tumours, tumours
CC of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer,
CC head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer,
CC gastric cancer, brain cancer, bladder cancer, prostate cancer and
CC urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours,
CC e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.
XX
XX Sequence 113 AA:
SQ
XX
XX Query Match 88.2%; Score 75; DB 21; Length 113;
XX Best Local Similarity 88.2%; Pred. No. 3.5e-05;
XX Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 KSSOSVLXSNHKNYLA 17
XX |||||||:|||||
XX Db 24 KSSOSVLXSNHKNYLA 40
XX
XX RESULT 14
XX AAB18873
XX ID AAB18873 standard; Protein; 113 AA.
XX
XX AC AAB18873;
XX
XX DT 08-FEB-2001 (first entry)
XX
XX DE Amino acid sequence of anti-p53 antibody light chain clone 163.17.
XX
XX KM p53; antibody; immune response; vaccine; gene therapy; cancer;
XX rheumatoid arthritis; coronary heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200056770-A1.
XX
XX PD 28-SEP-2000.
XX
XX PF 15-MAR-2000; 2000MO-AU00189.
XX
XX PR 19-MAR-1999; 99AU-0009321.
XX
XX PA (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX
XX Ward RL, Coomber DMJ;
XX
XX WPI; 2000-638249/61.
XX DR N-PSDB; AAA96150.
XX
XX PT Polynucleotides encoding anti-p53 antibodies, polypeptides and peptide
XX fragments, useful in treatment and diagnosis of cancer, Rheumatoid
XX arthritis and coronary heart disease
XX
XX PS Claim 30; Page 148; 163pp; English.
XX
XX The present sequence represents the light chain of an antibody reactive
XX against p53. The antibody is obtained from a vertebrate host expressing
XX an immune response against a naturally occurring disease. The antibodies

```

are useful in pharmaceutical compositions, which additionally contain chelators, drugs, prodrugs, toxins and imaging markers e.g. radioisotopes or gadolinium. The polypeptide components of the antibodies are useful in vaccines, for inducing an immune response against a disease in a vertebrate, for treatment and/or prophylaxis of disease and for detection purposes. The nucleic acid sequences can be used to detect a disease as well as for gene therapy and recombinant production of the polypeptides. In particular, the following can be treated cancer: rheumatoid arthritis and coronary heart disease. Cancers include carcinogenic tumours, tumours of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer, head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer, gastric cancer, brain cancer, bladder cancer, prostate cancer and urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours, e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.

Sequence 113 AA:

Query Match 88.2%; Score 75; DB 21; Length 113;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLA 17
|||||||:|:|:|
DB 24 KSSQSVLYSSNNKNYLA 40

RESULT 15

AAB18881

ID AAB18881 standard: Protein; 113 AA.

AC AAB18881;

DT 08-FEB-2001 (first entry)

DE Amino acid sequence of anti-p53 antibody light chain clone 163.24.

KM p53; antibody; immune response; vaccine; gene therapy; cancer;

OS Homo sapiens.

PN WO200056770-A1.

PD 28-SEP-2000.

PF 15-MAR-2000; 2000WO-AU00189.

PR 19-MAR-1999; 99AU-0009321.

PA (SVIN-) ST VINCENT'S HOSPITAL, SYDNEY LTD.

PI Ward RL, Coomber DWJ;

DR WPI: 2000-638249/61.

DR N-PSDB; AAA96158.

PS Claim 30; Page 156; 163pp; English.

The present sequence represents the light chain of an antibody reactive against p53. The antibody is obtained from a vertebrate host expressing an immune response against a naturally occurring disease. The antibodies are useful in pharmaceutical compositions, which additionally contain chelators, drugs, prodrugs, toxins and imaging markers e.g. radioisotopes or gadolinium. The polypeptide components of the antibodies are useful in vaccines, for inducing an immune response against a disease in a vertebrate, for treatment and/or prophylaxis of disease and for detection purposes. The nucleic acid sequences can be used to detect a disease as well as for gene therapy and recombinant production of the polypeptides. In particular, the following can be treated cancer, rheumatoid arthritis

and coronary heart disease. Cancers include carcinogenic tumours, tumours of epithelial origin, e.g. colo-rectal cancer, breast cancer lung cancer, head and neck tumours, hepatic cancer, pancreatic cancer, ovarian cancer, gastric cancer, brain cancer, bladder cancer, prostate cancer and urinary/genital tract cancer, oesophageal cancer, mesenchymal tumours, e.g. sarcoma, and hemopoietic tumours, e.g. B cell lymphoma.

Sequence 113 AA:

Query Match 88.2%; Score 75; DB 21; Length 113;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLA 17
|||||||:|:|:|
DB 24 KSSQSVLYSSNNKNYLA 40

Search completed: October 7, 2003, 19:13:52
Job time : 22.3966 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 8.7931 Seconds
(without alignments)
305,878 Million cell updates/sec

Title: US-09-988-013a-2_COPY_24_40
Perfect score: 85
Sequence: 1 KSSQSVLYSANHKNYLA 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	85	100.0	113	10	US-09-741-843-2
2	85	100.0	113	10	US-09-741-843-2
3	85	100.0	113	11	US-09-894-839-2
4	85	100.0	113	11	US-09-894-839-2
5	85	100.0	113	11	US-09-894-839-20
6	85	100.0	113	11	US-09-894-839-20
7	85	100.0	113	11	US-09-894-839-20
8	85	100.0	113	11	US-09-894-839-20
9	85	100.0	113	11	US-09-894-839-20
10	85	100.0	113	11	US-09-894-839-20
11	85	100.0	113	11	US-09-894-839-20
12	85	100.0	113	11	US-09-894-839-20
13	85	100.0	113	11	US-09-894-839-20
14	85	100.0	113	11	US-09-894-839-20
15	85	100.0	113	11	US-09-894-839-20
16	85	100.0	113	11	US-09-894-839-20
17	85	100.0	113	11	US-09-894-839-20
18	85	100.0	113	11	US-09-894-839-20

16	75	88.2	260	12	US-10-255-478-64	Sequence 64, Appl
17	75	88.2	274	12	US-10-255-478-66	Sequence 66, Appl
18	75	88.2	284	12	US-10-255-478-70	Sequence 70, Appl
19	74	87.1	112	12	US-10-229-335-4	Sequence 4, Appl
20	74	87.1	112	12	US-10-229-335-28	Sequence 28, Appl
21	74	87.1	112	15	US-10-056-052-6	Sequence 6, Appl
22	74	87.1	112	15	US-10-056-052-10	Sequence 10, Appl
23	74	87.1	112	15	US-10-056-052-18	Sequence 18, Appl
24	73	85.9	113	10	US-09-274-163E-16	Sequence 16, Appl
25	73	85.9	114	9	US-09-810-502-38	Sequence 38, Appl
26	73	85.9	114	10	US-09-274-163E-2	Sequence 2, Appl
27	73	85.9	114	10	US-09-274-163E-4	Sequence 4, Appl
28	73	85.9	114	10	US-09-274-163E-6	Sequence 6, Appl
29	73	85.9	114	10	US-09-274-163E-11	Sequence 11, Appl
30	73	85.9	114	10	US-09-274-163E-16	Sequence 16, Appl
31	73	85.9	114	10	US-09-274-163E-21	Sequence 21, Appl
32	72	84.7	94	9	US-09-905-243-60	Sequence 60, Appl
33	72	84.7	114	14	US-10-025-687-11	Sequence 11, Appl
34	72	84.7	114	15	US-10-125-687-11	Sequence 11, Appl
35	72	84.7	114	15	US-10-171-452A-1	Sequence 1, Appl
36	72	84.7	114	12	US-10-390-986-16	Sequence 16, Appl
37	71	83.5	112	15	US-10-056-052-14	Sequence 14, Appl
38	71	83.5	154	9	US-09-925-299-1226	Sequence 1226, Ap
39	71	83.5	154	11	US-09-925-299-1226	Sequence 1226, Ap
40	70	82.4	17	14	US-10-146-305-9	Sequence 9, Appl
41	70	82.4	113	10	US-09-999-025-7	Sequence 7, Appl
42	70	82.4	113	10	US-09-999-025-9	Sequence 9, Appl
43	70	82.4	113	10	US-09-999-025-13	Sequence 13, Appl
44	70	82.4	113	10	US-09-999-025-14	Sequence 14, Appl
45	70	82.4	113	10	US-09-999-040-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-741-843-2
Sequence 2, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
TITLE OF INVENTION: HANSEN, Hans
FILE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
CURRENT APPLICATION NUMBER: US/09/741,843
CURRENT FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: US 09/127,902
PRIORITY FILING DATE: 1998-08-03
PRIORITY APPLICATION NUMBER: US 08/690,102
PRIORITY FILING DATE: 1996-07-06
PRIORITY APPLICATION NUMBER: US 08/289,576
PRIORITY FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patent version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-741-843-2

Query Match 100.0%; Score 85; DB 10; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
DB 24 KSSQSVLYSANHKNYLA 40

RESULT 2
US-09-741-843-6
Sequence 6, Application US/09741843

Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LY
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-6

Query Match 100.0%; Score 85; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSSQSVLYSANHKNTLA 17
Db 24 KSSQSVLYSANHKNTLA 40

RESULT 3
US-09-894-839-2
; Sequence 2, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: LEUNG, Shui-on
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-2

Query Match 100.0%; Score 85; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSSQSVLYSANHKNTLA 17
Db 24 KSSQSVLYSANHKNTLA 40

RESULT 4
US-09-894-839-6
; Sequence 6, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing

; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-6

Query Match 100.0%; Score 85; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSSQSVLYSANHKNTLA 17
Db 24 KSSQSVLYSANHKNTLA 40

RESULT 5
US-09-894-839-20
; Sequence 20, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-20

Query Match 100.0%; Score 85; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSSQSVLYSANHKNTLA 17
Db 24 KSSQSVLYSANHKNTLA 40

RESULT 6
US-09-988-013a-2
; Sequence 2, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902

```
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-988-013a-2

Query Match
Best Local Similarity 100.0%; Score 85; DB 11; Length 113;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
Db 24 KSSQSVLYSANHKNYLA 40

RESULT 7
US-09-988-013a-6
; Sequence 6, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; TITLE OF INVENTION: CELL LYMPHOMA AND LEUKEMIA CELLS
; FILE REFERENCE: 18/33/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-988-013a-6

Query Match
Best Local Similarity 100.0%; Score 85; DB 11; Length 113;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
Db 24 KSSQSVLYSANHKNYLA 40

RESULT 8
US-09-563-222-14
; Sequence 14, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hiatt, Andrew
; APPLICANT: Heilm, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; TITLE OF INVENTION: EUKARYOTIC CELLS
; FILE REFERENCE: 310098.406
; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 14
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-563-222-14

Query Match
Best Local Similarity 88.2%; Score 75; DB 11; Length 17;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
Db 1 KSSQSVLYSSNNKNKYLA 17

RESULT 9
US-10-078-958-15
; Sequence 15, Application US/10078958
; Publication No. US20030070185A1
; GENERAL INFORMATION:
; APPLICANT: JAKOBOVITS, AYA
; APPLICANT: KUCHERLAPAT, RAJU
; APPLICANT: KLAPHOLT, SUSAN
; APPLICANT: MENDEZ, MICHAEL J.
; APPLICANT: GREEN, LARRY
; TITLE OF INVENTION: TRANSGENIC MAMMALS HAVING HUMAN Ig LOCI INCLUDING
; TITLE OF INVENTION: PLURAL VH AND VK REGIONS AND ANTIBODIES PRODUCED
; FILE REFERENCE: CELL 4.18 CON
; CURRENT APPLICATION NUMBER: US/10/078,958
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 08/759,620
; PRIOR FILING DATE: 1996-12-03
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-078-958-15

Query Match
Best Local Similarity 88.2%; Score 75; DB 15; Length 82;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
Db 5 KSSQSVLYSSNNKNKYLA 21

RESULT 10
US-10-025-687-25
; Sequence 25, Application US/10025687
; Publication No. US20020142255A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Peter
; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
; FILE REFERENCE: 26050-705
; CURRENT APPLICATION NUMBER: US/10/025,687
; CURRENT FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-687-25

Query Match
Best Local Similarity 88.2%; Score 75; DB 14; Length 101;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNYLA 17
```

Db 24 KSSQSVLYSSNNKNYLA 40

RESULT 11
US-10-194-975-89
Sequence 89, Application US/10194975
Publication No. US20030039649A1
GENERAL INFORMATION:
APPLICANT: Foote, Jefferson
TITLE OF INVENTION: Super Humanized Antibodies
FILE REFERENCE: 501231.01
CURRENT APPLICATION NUMBER: US/10/194,975
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: US 60/305,111
PRIOR FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn version 3.1
SEQ ID NO 89
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapiens
US-10-194-975-89

Query Match 88.2%; Score 75; DB 15; Length 101;
Best Local Similarity 88.2%; Pred. No. 6.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSSNNKNYLA 17
Db 24 KSSQSVLYSSNNKNYLA 40

RESULT 12
US-10-125-687-25
Sequence 25, Application US/10125687
Publication No. US20030054407A1
GENERAL INFORMATION:
APPLICANT: Luo, Peter
TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
FILE REFERENCE: 26050-705
CURRENT APPLICATION NUMBER: US/10/125,687
CURRENT FILING DATE: 2002-04-17
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.1
SEQ ID NO 25
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-687-25

Query Match 88.2%; Score 75; DB 15; Length 101;
Best Local Similarity 88.2%; Pred. No. 6.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSSNNKNYLA 17
Db 24 KSSQSVLYSSNNKNYLA 40

RESULT 13
US-09-956-206A-80
Sequence 80, Application US/09956206A
Patent No. US20020164339A1
GENERAL INFORMATION:
APPLICANT: DO COUTO, FERNANDO J. R.
CERIANI, ROBERTO L.
PETERSON, JERRY A.
TITLE OF INVENTION: RECOMBINANT PEPTIDES DERIVED FROM THE
MC3 ANTI-BA6 ANTIBODY, METHODS OF USE THEREOF, AND
METHODS OF HUMANIZING ANTIBODY PEPTIDES
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/956,206A
FILING DATE: 19-Apr-2002
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,539
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: PCT/US95/11683
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: 08/487,598
FILING DATE: 7-JUNE-1995
APPLICATION NUMBER: 08/307,868
FILING DATE: 16-SEPT-1994

ATTORNEY/AGENT INFORMATION:
NAME: WITT, ERIC
REGISTRATION NUMBER: 44,408
REFERENCE/DOCKET NUMBER: 276332000101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-956-206A-80

Query Match 88.2%; Score 75; DB 10; Length 113;
Best Local Similarity 88.2%; Pred. No. 7.1e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSSNNKNYLA 17
Db 24 KSSQSVLYSSNNKNYLA 40

RESULT 14
US-09-995-529-6
Sequence 6, Application US/09995529
Publication No. US2003009655A1
GENERAL INFORMATION:
APPLICANT: Watkins, Jeffrey D.
Huse, William D.
TITLE OF INVENTION: Humanized Collagen Antibodies and
FILE REFERENCE: P-IX 4976
CURRENT APPLICATION NUMBER: US/09/995,529
CURRENT FILING DATE: 2001-11-26
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-09-995-529-6

Query Match 88.2%; Score 75; DB 11; Length 113;
Best Local Similarity 88.2%; Pred. No. 7.1e-05;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLA 17
|||||||:|:|:|
DB 24 KSSQSVLYSSNNKNTLA 40

RESULT 15
US-10-255-478-58

; Sequence 58, Application US/10255478
; Publication No. US20030165498A1

; GENERAL INFORMATION:

; APPLICANT: Mezes, Peter S.

; APPLICANT: Richard, Ruth A.

; APPLICANT: Johnson, Kimberly S.

; APPLICANT: Schiom, Jeffrey

; APPLICANT: Kashmiri, Syed V.S.

; APPLICANT: Shu, Liming

; APPLICANT: Padian, Eduardo A.

; TITLE OF INVENTION: Composite Antibodies of Humanized Human Subgroup IV Light Chain

; FILE REFERENCE: 37777E

; CURRENT APPLICATION NUMBER: US/10/255,478

; PRIOR APPLICATION NUMBER: 2002-09-25

; PRIOR FILING DATE: 1997-10-30

; PRIOR APPLICATION NUMBER: US 60/030,173

; PRIOR FILING DATE: 1996-10-31

; PRIOR APPLICATION NUMBER: US 08/261,354

; PRIOR FILING DATE: 1994-06-16

; PRIOR APPLICATION NUMBER: US 07/964,536

; PRIOR FILING DATE: 1992-10-20

; PRIOR APPLICATION NUMBER: US 07/510,697

; PRIOR FILING DATE: 1990-07-17

; NUMBER OF SEQ ID NOS: 78

; SOFTWARE: Microsoft Word 97 SR-2

; SEQ ID NO 58

; LENGTH: 134

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: Hum4 VL

; LOCATION: 1..134

; US-10-255-478-58

Query Match 88.2%; Score 75; DB 12; Length 134;

Best Local Similarity 88.2%; Pred. No. 8.5e-05;

Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSQSVLYSANHKNTLA 17
|||||||:|:|:|

DB 44 KSSQSVLYSSNNKNTLA 60

Search completed: October 7, 2003, 19:24:24

Job time : 9.7931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 6.33103 Seconds
(without alignments)
258.231 million cell updates/sec

Title: US-09-988-013a-2_COPY_24_40
Perfect score: 85
Sequence: 1 KSSOSVLYSANHNKNTLA 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	88.2	92	2	Ig kappa chain V r
2	75	88.2	92	2	Ig kappa chain V r
3	75	88.2	92	2	Ig kappa chain V r
4	75	88.2	92	2	Ig kappa chain V r
5	75	88.2	92	2	Ig kappa chain V r
6	75	88.2	92	2	Ig kappa chain V r
7	75	88.2	92	2	Ig kappa chain V r
8	75	88.2	106	2	Ig kappa chain V r
9	75	88.2	113	2	Ig kappa chain V r
10	75	88.2	113	2	Ig kappa chain V r
11	75	88.2	120	2	Ig kappa chain V r
12	75	88.2	121	1	antibody light cha
13	75	88.2	133	1	Ig kappa chain pre
14	75	88.2	134	1	Ig kappa chain pre
15	75	88.2	134	1	anti-sm antibody V
16	74	87.1	103	2	Ig light chain V r
17	74	87.1	103	2	Ig light chain V r
18	73	85.9	114	1	Ig kappa chain V r
19	72	84.7	114	1	Ig kappa chain V-I
20	72	84.7	129	2	Ig kappa chain V-I
21	71	83.5	101	2	Ig kappa chain - h
22	71	83.5	101	2	Ig kappa chain V r
23	71	83.5	112	2	Ig kappa chain V r
24	70	82.4	113	2	Ig kappa chain V r
25	69	81.2	92	2	antitumor monoclon
26	69	81.2	92	2	Ig kappa chain V r
27	68	80.0	134	1	Ig kappa chain pre
28	68	80.0	96	2	Ig kappa chain V r
29	68	80.0	112	2	Ig kappa chain V r
			113	2	Ig kappa chain V r

30	68	80.0	138	2	S26040	Ig kappa chain pre
31	67	78.8	124	2	S40364	Ig kappa chain - h
32	66	77.6	114	2	S44116	Ig kappa chain V-J
33	66	77.6	132	2	S46373	Ig kappa chain V-J
34	66	77.6	240	2	S06084	Ig kappa chain pre
35	65	76.5	102	2	A34153	Ig kappa chain V-I
36	62	72.9	104	2	PH1053	Ig kappa chain V r
37	62	72.9	113	2	JC2270	Ig light chain V r
38	62	72.9	134	2	PC1214	Ig kappa chain pre
39	61	71.8	105	2	C30535	Ig kappa chain V r
40	61	71.8	107	2	G30535	Ig kappa chain V r
41	61	71.8	107	2	D30535	Ig kappa chain V r
42	61	71.8	107	2	F30535	Ig kappa chain V r
43	61	71.8	108	2	E30535	Ig kappa chain V r
44	61	71.8	112	2	E30538	Ig kappa chain V r
45	61	71.8	112	2	F30538	Ig kappa chain V r

ALIGNMENTS

RESULT 1
S37533
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37533
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IgD(+) cells, the major B cell subset in the peripheral bl
A:Reference number: S37501
A:Accession: S37533
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:226627; NID:g405708; PIDN:CA81380.1; PID:g405709
A:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin

Query Match 88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHNKNTLA 17
DB 2 KSSOSVLYSSNNKNTLA 18

RESULT 2

S37530
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37530
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IgD(+) cells, the major B cell subset in the peripheral bl
A:Reference number: S37501
A:Accession: S37530
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:226630; NID:g405702; PIDN:CA81383.1; PID:g405703
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin

Query Match 88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHNKNTLA 17
DB 2 KSSOSVLYSSNNKNTLA 18

```
RESULT 3
S37529
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37529
R:Klein, U.; Kupepers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IGD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26631; NID:9405700; PIDN:CAA81384.1; PID:9405701
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 KSSOSVLYSANHKNYLA 17
        |||||||:|:|||||
Db      2 KSSOSVLYSSNNKNYLA 18

RESULT 4
S37535
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37535
R:Klein, U.; Kupepers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IGD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37535
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26625; NID:9405712; PIDN:CAA81378.1; PID:9405713
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 KSSOSVLYSANHKNYLA 17
        |||||||:|:|||||
Db      2 KSSOSVLYSSNNKNYLA 18

RESULT 5
S37534
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37534
R:Klein, U.; Kupepers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IGD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37534
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26626; NID:9405710; PIDN:CAA81379.1; PID:9405711
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
```

```
Query Match      88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 KSSOSVLYSANHKNYLA 17
        |||||||:|:|||||
Db      2 KSSOSVLYSSNNKNYLA 18

RESULT 6
S37532
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37532
R:Klein, U.; Kupepers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+) IGD(+) cells, the major B cell subset in the peripheral bl
A:Reference number: S37501
A:Accession: S37532
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26628; NID:9405706; PIDN:CAA81381.1; PID:9405707
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      88.2%; Score 75; DB 2; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 KSSOSVLYSANHKNYLA 17
        |||||||:|:|||||
Db      2 KSSOSVLYSSNNKNYLA 18

RESULT 7
A49138
IGA kappa rheumatoid factor variable - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
C:Accession: A49138
R:Gause, A.; Kuipers, R.; Mierau, R.
Clin. Exp. Immunol. 88, 430-434, 1992
A:Title: A somatically mutated V kappa IV gene encoding a human rheumatoid factor 11g
A:Reference number: A49138; MUID:92298590; PMID:1606727
A:Accession: A49138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <GAU>
A:Cross-references: GB:S37926; NID:9298207; PIDN:AAB22366.1; PID:9298208
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCI backbone (NCIN:106633, NCBI:P106637)
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:20-99/Domain: immunoglobulin homology <IM>

Query Match      88.2%; Score 75; DB 2; Length 106;
Best Local Similarity 88.2%; Pred. No. 1.5e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 KSSOSVLYSANHKNYLA 17
        |||||||:|:|||||
Db      28 KSSOSVLYSSNNKNYLA 44

RESULT 8
S30523
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S30523
R:Marlette, X.
```


submitted to the EMBL Data Library, October 1992
A:Reference number: S30520

A:Accession: S30523
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <MAR>
A:Cross-references: EMBL:Z18329
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 88.2%; Score 75; DB 2; Length 113;
Best Local Similarity 88.2%; Pred. No. 1.6e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKNTLA 17
|||||||:|:|||||
DB 24 KSSOSVLYSSNNKNTLA 40

RESULT 9

S34003
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S34003
R:Marlette, X.; Tsapis, A.; Brouet, J.C.
Eur. J. Immunol. 23, 846-851, 1993
A:Title: Nucleotide sequence analysis of the variable domains of four human monoclonal
A:Reference number: S34001; MUID:93209281; PMID:7681398
A:Accession: S34003
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <MAR>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 88.2%; Score 75; DB 2; Length 113;
Best Local Similarity 88.2%; Pred. No. 1.6e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKNTLA 17
|||||||:|:|||||
DB 24 KSSOSVLYSSNNKNTLA 40

RESULT 10

S34002
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S34002; S30522
R:Marlette, X.; Tsapis, A.; Brouet, J.C.
Eur. J. Immunol. 23, 846-851, 1993
A:Title: Nucleotide sequence analysis of the variable domains of four human monoclonal
A:Reference number: S34001; MUID:93209281; PMID:7681398
A:Accession: S34002
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <MAR>
A:Cross-references: EMBL:Z18328
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 88.2%; Score 75; DB 2; Length 113;
Best Local Similarity 88.2%; Pred. No. 1.6e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKNTLA 17
|||||||:|:|||||
DB 24 KSSOSVLYSSNNKNTLA 40

RESULT 11

S51147
antibody light chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000
C:Accession: S51147
R:de Kruijf, J.; Boel, E.; Logtenberg, T.
submitted to the EMBL Data Library, January 1995
A:Description: Selection and application of human SCFV antibody fragments from a semi
A:Reference number: S51147
A:Accession: S51147
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <DEK>
A:Cross-references: EMBL:X83714; NID:96333227; PIDN:CA58689.1; PID:96333228
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 88.2%; Score 75; DB 2; Length 120;
Best Local Similarity 88.2%; Pred. No. 1.7e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKNTLA 17
|||||||:|:|||||
DB 24 KSSOSVLYSSNNKNTLA 40

RESULT 12

K4H0
Ig kappa chain precursor V-IV region - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
C:Accession: A01902
R:Klobeck, H.G.; Bornkamm, G.W.; Combrato, G.; Mocikat, R.; Pohlenz, H.D.; Zachau, H.
Nucleic Acids Res. 13, 6515-6529, 1985
A:Title: Subgroup IV of human immunoglobulin K light chains is encoded by a single ge
A:Reference number: A93589; MUID:86041853; PMID:2997712
A:Accession: A01902
A:Molecule type: DNA
A:Residues: 1-121 <KLO>
A:Note: the sequence was determined from the germline gene
A:Note: there is only one Ig kappa V-IV gene
C:Genetics:

A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as Iga and Igh, the subunits associate into
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-121/Product: Ig kappa chain precursor V-IV region #status predicted <MAT>
F:21-43/Region: framework 1
F:36-116/Domain: immunoglobulin homology <IMM>
F:44-60/Region: complementarity-determining 1
F:61-75/Region: complementarity-determining 2
F:76-82/Region: complementarity-determining 3
F:83-114/Region: framework 3
F:115-121/Region: complementarity-determining 3
F:43-114/Disulfide bonds: #status predicted

Query Match 88.2%; Score 75; DB 1; Length 121;
Best Local Similarity 88.2%; Pred. No. 1.7e-05;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKNTLA 17
|||||||:|:|||||
DB 44 KSSOSVLYSSNNKNTLA 60

RESULT 13

K4H0J1
Ig kappa chain precursor V-IV region (J1) - human

C/Species: Homo sapiens (man)
 C/Date: 30-Jun-1987 #sequence-revision 30-Jun-1987 #text_change 21-Jan-2000
 C/Accession: A01904
 R:Klobeck, H.G.; Bornkamm, G.W.; Combrinato, G.; Moolkat, R.; Pohlentz, H.D.; Zachau, H.G.
 Nucleic Acids Res. 13, 6515-6529, 1985
 A/Title: Subgroup IV of human immunoglobulin K light chains is encoded by a single germ
 A/Reference number: A93889; MID:86041853; PMID:297712
 A/Accession: A01904
 A/Molecule type: DNA
 A/Residues: 1-133 <RIO>
 A/Cross-references: GB:200022; GB:X51570; NID:933158; PIDN:CA477317.1; PID:g296654
 A/Note: the sequence was determined from the differentiated gene
 C/Genetics:
 A/Gene: GDB:IGKV
 A/Cross-references: GDB:119341; ONIM:146980
 A/Map position: 2p12-2p12
 A/Intons: 17/1
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C/Keywords: heterotetramer; Immunoglobulin
 F/1-20/Domain: signal sequence #status predicted <SIG>
 F/21-133/Product: Ig kappa chain V-IV region (JI) #status predicted <MAT>
 F/21-43/Region: framework 1
 F/36-116/Domain: immunoglobulin homology <IMM>
 F/44-60/Region: complementarity-determining 1
 F/61-75/Region: framework 2
 F/76-82/Region: complementarity-determining 2
 F/83-114/Region: framework 3
 F/115-122/Region: complementarity-determining 3
 F/123-133/Region: framework 4
 F/43-114/Disulfide bonds: #status predicted

Query Match 88.2%; Score 75; DB 1; Length 133;
 Best Local Similarity 88.2%; Pred. No. 1.9e-05;
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSSNHNKNYLA 17
 DB 44 KSSOSVLYSSNHNKNYLA 60

RESULT 14
 S49531
 anti-Sm antibody VL chain (V kappa 4/J kappa 3) - human
 C/Species: Homo sapiens (man)
 C/Date: 01-Feb-1995 #sequence-revision 12-May-1995 #text_change 21-Jan-2000
 C/Accession: S49531
 R:Mahmoudi, M.; Edwards, J.; Cairns, E.; Bell, D.
 submitted to the EMBL Data Library, October 1994
 A/Description: Molecular characterization of natural human anti-Sm autoantibodies.
 A/Reference number: S48797
 A/Accession: S49531
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-134 <MAH>
 A/Cross-references: EMBL:246347; NID:g560841; PIDN:CA486466.1; PID:g560842
 C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
 F/36-116/Domain: immunoglobulin homology <IMM>

Query Match 88.2%; Score 75; DB 2; Length 134;
 Best Local Similarity 88.2%; Pred. No. 1.9e-05;
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSSNHNKNYLA 17
 DB 44 KSSOSVLYSSNHNKNYLA 60

RESULT 15
 S26337
 Ig light chain V region - mouse (fragment)
 C/Species: Mus musculus (house mouse)

C/Date: 19-Mar-1998 #sequence-revision 19-Mar-1998 #text_change 20-Jun-2000
 C/Accession: S26337; S78449
 R:Stark, S.E.; Caton, A.J.
 J. Exp. Med. 174, 613-624, 1991
 A/Title: Antibodies that are specific for a single amino acid interchange in a protei
 A/Reference number: S26309; MID:91341421; PMID:1908510
 A/Accession: S26337
 A/Molecule type: mRNA
 A/Residues: 1-101 <STA>
 A/Cross-references: EMBL:X59193
 R:Caton, A.J.
 submitted to the EMBL Data Library, April 1991
 A/Reference number: S78447
 A/Accession: S78449
 A/Molecule type: mRNA
 A/Residues: 1-60, 'T', 'G', 'S', '93-101 <CAT>
 A/Cross-references: EMBL:X59193; NID:g52323; PIDN:CA441903.1; PID:g1334067
 C/Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C/Keywords: heterotetramer; Immunoglobulin
 F/8-88/Domain: immunoglobulin homology <IMM>

Query Match 87.1%; Score 74; DB 2; Length 101;
 Best Local Similarity 88.2%; Pred. No. 2.1e-05;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KSSOSVLYSSNHNKNYLA 17
 DB 16 KSSOSVLYSSNHNKNYLA 32

Search completed: October 7, 2003, 19:21:47
 Job time : 6.33103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 3.45862 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-2_COPY_24_40

Sequence: 1 KSSQSVLYSNHKNYIA 17

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Swissprot_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	88.2	121	1	P06312 homo sapien
2	75	88.2	133	1	P06313 homo sapien
3	73	85.9	114	1	P01625 homo sapien
4	72	84.7	109	1	P83593 homo sapien
5	69	81.2	134	1	P06314 homo sapien
6	42	49.4	1529	1	P04182 saccharomyc
7	41	48.2	1237	1	P53254 saccharomyc
8	40	47.1	365	1	P94168 mus musculi
9	40	47.1	517	1	O99446 ctenorhadi
10	39	45.9	364	1	O99446 ctenorhadi
11	39	45.9	680	1	O99446 ctenorhadi
12	39	45.9	770	1	O99446 ctenorhadi
13	39	45.9	773	1	O99446 ctenorhadi
14	38.5	45.3	117	1	O99446 ctenorhadi
15	38	44.7	214	1	O99446 ctenorhadi
16	38	44.7	331	1	O99446 ctenorhadi
17	38	44.7	337	1	O99446 ctenorhadi
18	38	44.7	337	1	O99446 ctenorhadi
19	38	44.7	388	1	O99446 ctenorhadi
20	38	44.7	411	1	O99446 ctenorhadi
21	38	44.7	525	1	O99446 ctenorhadi
22	38	44.7	610	1	O99446 ctenorhadi
23	38	44.7	1045	1	O99446 ctenorhadi
24	38	44.7	1140	1	O99446 ctenorhadi
25	38	44.7	1140	1	O99446 ctenorhadi
26	38	44.7	3649	1	O99446 ctenorhadi
27	37.5	44.1	1941	1	O99446 ctenorhadi
28	37	43.5	239	1	O99446 ctenorhadi
29	37	43.5	252	1	O99446 ctenorhadi
30	37	43.5	361	1	O99446 ctenorhadi
31	37	43.5	361	1	O99446 ctenorhadi
32	37	43.5	430	1	O99446 ctenorhadi
33	37	43.5	462	1	O99446 ctenorhadi

34	37	43.5	569	1	SYE_TOBAC	Q43794 nicotiana t
35	37	43.5	601	1	T248_CHLMO	Q9P160 chlamydia m
36	37	43.5	725	1	VGLH_MCVWS	P30673 murine cyto
37	37	43.5	858	1	ALR2_YEAST	P43533 saccharomyc
38	37	43.5	859	1	ALR1_YEAST	O08269 saccharomyc
39	37	43.5	1111	1	KRPL_YEAST	P28742 saccharomyc
40	37	43.5	1117	1	TERF_YEAST	O77448 tetrahymena
41	37	43.5	1345	1	VCAP_HSV7J	P52347 human herpe
42	37	43.5	6669	1	NEBU_HUMAN	P20929 homo sapien
43	36	42.4	196	1	HIS5_BUCBP	P59501 buchnera ap
44	36	42.4	221	1	MLTE_BUCAI	P57352 buchnera ap
45	36	42.4	284	1	MODD_ECO57	P58496 escherichia

ALIGNMENTS

RESULT 1	ID	Sequence	Standard	PRT	121 AA
KV40_HUMAN	KV40_HUMAN	STANDARD	PRT	121 AA	
AC	P06312				
DT	01-JAN-1988 (Rel. 06, Created)				
DT	01-JAN-1988 (Rel. 06, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Ig kappa chain V-IV region precursor (Fragment).				
GN	IGKV4-1.				
OS	Homo sapiens (Human).				
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEBLIND=86041853; PubMed=2997712;				
RA	Klobeck H.G., Bornkamm G.W., Combrato R., Mochkat R., Pohlenz H.D.,				
RA	Zachau H.G.;				
RT	"Subgroup IV of human immunoglobulin K light chains is encoded by a				
RL	single germline gene.";				
CC	Nucleic Acids Res. 13:6515-6529(1985).				
CC	-1- MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; Z00023; CAA7318.1; -				
DR	PIR; A01902; K4H0.				
DR	HSSP; P80362; 1WTL.				
DR	Genew; HGNC:5834; IGKV4-1.				
DR	GO; GO:0005576; C:extracellular; NAS.				
DR	GO; GO:0003823; P:antigen binding activity; NAS.				
DR	GO; GO:0006955; P:immune response; NAS.				
DR	InterPro; IPR007110; Ig-like.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR003596; Ig_V.				
DR	Pfam; PF00047; Ig_1.				
DR	SMART; SM00406; Igv_1.				
DR	PROSITE; PS50835; IG-LIKE; 1.				
DR	Immunoglobulin V region; Signal.				
FW	SIGNAL				
FT	CHAIN	21	>121		IG KAPPA CHAIN V-IV REGION.
FT	DOMAIN	21	43		FRAMEWORK-1.
FT	DOMAIN	44	60		COMPLEMENTARITY-DETERMINING-1.
FT	DOMAIN	61	75		FRAMEWORK-2.
FT	DOMAIN	76	82		COMPLEMENTARITY-DETERMINING-2.
FT	DOMAIN	83	114		FRAMEWORK-3.
FT	DOMAIN	115	121		COMPLEMENTARITY-DETERMINING-3.
FT	DISULFID	43	114		BY SIMILARITY.
FT	NON_TER	121	121		
SO	SEQUENCE	121 AA;	13380 MW;	9586AD418BD33974	CRC64;

Query Match 88.2%; Score 75; DB 1; Length 121;
 Best Local Similarity 88.2%; Pred. No. 2.7e-06;
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANKNYLA 17
 |||||:|||||
 DB 44 KSSOSVLYSSNNKNYLA 60

RESULT 2

KV4B_HUMAN STANDARD; PRT; 133 AA.
 AC P06313;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-IV region JI precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-86041853; PubMed-2997712;
 RA Klobbeck H.G., Borkham G.W., Combrato G., Mockat R., Pohlentz H.D.,
 RA Zachau H.G.;
 RT Supergroup IV of human immunoglobulin K light chains is encoded by a
 RT single germline gene.
 RL Nucleic Acids Res. 13:6515-6529(1985).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL_200022; CAA77317.1; -
 DR PIR; A01904; K4H0J1.
 DR HSSP; P80362; 1MTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; P:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 133 IG KAPPA CHAIN V-IV REGION JI.
 FT DOMAIN 21 43 FRAMEWORK-1.
 FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 61 75 FRAMEWORK-2.
 FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 83 114 FRAMEWORK-3.
 FT DOMAIN 115 122 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 123 132 FRAMEWORK-4.
 FT DISULFID 43 114 BY SIMILARITY.
 FT NON_TER 133
 SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;

Query Match 88.2%; Score 75; DB 1; Length 133;
 Best Local Similarity 88.2%; Pred. No. 3e-06;
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANKNYLA 17
 |||||:|||||
 DB 44 KSSOSVLYSSNNKNYLA 60

RESULT 3
 KV4A_HUMAN STANDARD; PRT; 114 AA.
 ID P01625;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig kappa chain V-IV region Len.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-76004342; PubMed-50995;
 RA Schneider M., Hillebrand N.;
 RT "The primary structure of a monoclonal immunoglobulin-L-chain of
 RT subgroup IV of the kappa type (Bence-Jones protein Len).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
 RN [2]
 RP REVISION TO 9.
 RA Salomon A.;
 RL Submitted (Aug-1996) to the SWISS-PROT data bank.

CC -1 MISCELLANEOUS; THIS IS A BENCE-JONES PROTEIN.
 DR PDB; 1EEQ; 01-FEB-01.
 DR PDB; 1EEU; 03-FEB-01.
 DR PDB; 1EFQ; 09-FEB-01.
 DR PDB; 1EK3; 06-MAR-01.
 DR PDB; 1IVE; 21-JAN-98.
 DR PDB; 3IVE; 18-MAY-99.
 DR PDB; 5IVE; 28-MAR-01.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; P:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region; Bence-Jones protein; 3D-structure.
 FT DOMAIN 1 23 FRAMEWORK-1.
 FT DOMAIN 24 40 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 41 55 FRAMEWORK-2.
 FT DOMAIN 56 62 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 63 94 FRAMEWORK-3.
 FT DOMAIN 95 101 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 102 113 FRAMEWORK-4.
 FT DISULFID 23 94 BY SIMILARITY.
 FT NON_TER 114
 SQ SEQUENCE 114 AA; 12640 MW; 0647FD17E236485 CRC64;

Query Match 85.9%; Score 73; DB 1; Length 114;
 Best Local Similarity 88.2%; Pred. No. 5.6e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANKNYLA 17
 |||||:|||||
 DB 24 KSSOSVLYSSNNKNYLA 40

RESULT 4

KV4D_HUMAN STANDARD; PRT; 109 AA.
 ID P83593;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig kappa chain V-IV region 5TH (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE.
 RC TISSUE=Abdominal adipose tissue;
 RX PubMed=9588180;
 RA Olsen K.E., Sletten K., Westermarck P.;
 RT "Extended analysis of AL-amyloid protein from abdominal wall
 subcutaneous fat biopsy: kappa IV immunoglobulin light chain.";
 RL Biochem. Biophys. Res. Commun. 245:713-716(1998).
 CC -1- FUNCTION: May play an important role in fibrillogenesis.
 DR InterPro: IPR003599; IG_1.
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; IG_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 23 FRAMEWORK-1.
 FT DOMAIN 24 40 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 41 55 FRAMEWORK-2.
 FT DOMAIN 56 62 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 63 94 FRAMEWORK-3.
 FT DOMAIN 95 101 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 102 109 FRAMEWORK-4.
 FT DISULFD 23 94 BY SIMILARITY.
 FT UNSURE 23 23
 FT UNSURE 94 94
 FT NON_TER 109 109
 SQ SEQUENCE 109 AA; 12060 MW; 0C4F31EAL1E12A0B CRC64;
 Query Match 84.7%; Score 72; DB 1; Length 109;
 Best Local Similarity 82.4%; Pred. No. 8.1e-06;
 Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KSSOSVLYSANHKNYLA 17
 Db 24 KSSOSVLYSSNKNKNTLA 40
 RESULT 5
 KVALC_HUMAN STANDARD; PRT; 134 AA.
 ID P06314;
 DT 01-JAN-1988 (Rel. 05, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-IV region B17 precursor.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86041854; PubMed=2997713;
 RA Marsh P., Mills F., Gould H.;
 RT "Detection of a unique human V kappa IV germline gene by a cloned
 cDNA probe.";
 RL Nucleic Acids Res. 13:6531-6544(1985).
 CC [2]
 RP REVISION TO 76.
 RA Marsh P.;
 RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: X02990; CAA26733.1; -
 DR HSSP: P80362; IWTL.

DR GO: 0005576; C:extracellular; NAS.
 DR GO: 0003823; F:antigen binding activity; NAS.
 DR GO: 0006955; P:immune response; NAS.
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; IG_1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 134 IG KAPPA CHAIN V-IV REGION B17.
 FT DOMAIN 21 43 FRAMEWORK-1.
 FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 61 75 FRAMEWORK-2.
 FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 83 114 FRAMEWORK-3.
 FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 122 133 FRAMEWORK-4.
 FT DISULFD 43 114 BY SIMILARITY.
 FT NON_TER 134 134
 SQ SEQUENCE 134 AA; 14966 MW; 6413A22FD0738832 CRC64;
 Query Match 81.2%; Score 69; DB 1; Length 134;
 Best Local Similarity 76.5%; Pred. No. 3.4e-05;
 Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KSSOSVLYSANHKNYLA 17
 Db 44 KSSOSVLYSSDNKNKNTLA 60
 RESULT 6
 PDRF_YEAST STANDARD; PRT; 1529 AA.
 ID PDRF_YEAST
 AC O04182;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ATP-dependent permease PDR15.
 GN PDR15 OR YDR406W OR D9509.24.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S., Milligan J., Allen E., Araujo R., Aviles E., Beino A.,
 RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunkeler-Smith S., Hyman R., Kamp C., LaShkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILAR: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: U32274; AAB64846.1; -
 DR PIR: S69686; S69688.
 DR SGD: S0002814; PDR15.
 DR InterPro: IPR003593; AAA_Artpase.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR005285; PDR.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.

DR SMART: SM00382; AAA; 2.
 DR TIGR00956; 3a01205; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR ATP-binding; Transmembrane; Glycoprotein; Transport.
 FT DOMAIN 1 531 POTENTIAL.
 FT TRANSMEM 532 552 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 566 588 POTENTIAL.
 FT TRANSMEM 618 638 POTENTIAL.
 FT TRANSMEM 643 663 POTENTIAL.
 FT TRANSMEM 700 720 POTENTIAL.
 FT TRANSMEM 784 804 POTENTIAL.
 FT DOMAIN 805 1219 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1220 1240 POTENTIAL.
 FT TRANSMEM 1313 1333 POTENTIAL.
 FT TRANSMEM 1341 1361 POTENTIAL.
 FT TRANSMEM 1369 1389 POTENTIAL.
 FT TRANSMEM 1397 1417 POTENTIAL.
 FT TRANSMEM 1493 1513 POTENTIAL.
 FT DOMAIN 1514 1529 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 920 927 ATP (POTENTIAL).
 FT DOMAIN 13 24 POLY-SER.
 FT DOMAIN 66 69 POLY-SER.
 FT DOMAIN 794 797 POLY-PHE.
 FT CARBOHYD 558 558 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 744 744 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1529 AA; 172255 MW; DB58FB2E0534974 CRC64;

Query Match 49.4%; Score 42; DB 1; Length 1529;
 Best Local Similarity 61.5%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSOVSXSNHKN 14
 DB 53 TSOSLYTANSNN 65

RESULT 7
 YG2L_YEAST STANDARD; PRT; 1237 AA.
 AC P53254;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hypothetical 140.5 kDa protein in CTR1-PRP31 intergenic region.
 GN YG000W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wedler H., Scharfe M., Wedler E., Wambutt R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hernandez K., Weber N., Wipfl P., Schmidheini T.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY TO S.POMBE SPBC776.08C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z72875; CA97093.1;
 DR PIR: S64385; S64385.
 DR SGD: S0003322; YG000W.
 DR InterPro: IPR005554; Nrap.
 DR Pfam: PF03813; Nrap; 1.

KW Hypothetical protein.
 SQ SEQUENCE 1237 AA; 140484 MW; 9A2B5C885493D7D3 CRC64;
 Query Match 48.2%; Score 41; DB 1; Length 1237;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SYLYSANHKNYL 16
 DB 335 SYLYSSTHENYL 346

RESULT 8
 PICE_MOUSE STANDARD; PRT; 365 AA.
 ID PICE_MOUSE
 AC 09DIE8; 08B61; 08CGN6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lygophosphatidic acid
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate
 DE O-acyltransferase 5).
 GN AGPAT5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI;
 RA Lu B., Jiang Y.J., Chan M., Choy P.C.;
 RT Identification and characterization of 1-acylglycerolphosphate
 RT acyltransferase-epsilon.
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J, and NOD; TISSUE=Cerebellum, Embryo, and Spleen;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamana H., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Balderelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanpin A., Matsuda H., Batalov S., Heisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochla C., Cordani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Seton M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verato R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loguanello N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT *Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.*
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
 CC -1- PATHWAY: De novo phospholipid biosynthesis, second step.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
 CC ACYLTRANSFERASE FAMILY.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-12 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AY161042; AAN75571.1; -
 DR EMBL: AK003649; BAB22915.2; -
 DR EMBL: AK082137; BAC38421.1; -
 DR EMBL: AK089885; BAC40983.1; -
 DR EMBL: BC031987; AAB31987.1; ALT-INIT.
 DR MGD: MGI:1915880; 1110013A05R1K.
 DR InterPro: IPR002123; Acyltransferase.
 DR Pfam: PF01553; Acyltransferase; 1.
 DR SMART: SM00563; PISC; 1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 KW Transmembrane.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 345 365 POTENTIAL.
 FT CONFLICT 80 80 L -> W (IN REF. 1).
 FT CONFLICT 128 128 L -> M (IN REF. 1).
 SO SEQUENCE 365 AA; 42202 MW; C13E14759610E19B CRC64.

Query Match 47.1%; Score 40; DB 1; Length 365;
 Best Local Similarity 42.9%; Pred. No. 14;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKN 14
 DB 82 KKEKENVITLNNHOS 95
 | : : : : | : : : :
 KSSOSVLYSANHKN 14
 KKEKENVITLNNHOS 95

RESULT 9
 SEST_CAEEL STANDARD: PRT: 517 AA.
 AC Q9NAD6; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Putative sestrin.
 GN Y7AC9A.5.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Du H., Wohlmann P., Ames M.,
 RL submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RL Axelsten K., Balroch A.;
 CC Unpublished observations (MAR-2001).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE sestrin family.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AC024206; AAF36051.1; ALT_SEQ.
 DR WormPep: Y7AC9A.5; CE24663.
 DR InterPro: IPR006730; PA26.
 DR Pfam: PF04636; PA26; 1.
 KW Hypothetical protein; Nuclear protein.
 SO SEQUENCE 517 AA; 60881 MW; D7BC041916DDE205 CRC64;

Query Match 47.1%; Score 40; DB 1; Length 517;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 SSSOSVLYSANHKN 15
 DB 303 SAPATYSSNNRKNY 316
 | : : : : | : : : :
 SSSOSVLYSANHKN 15
 SAPATYSSNNRKNY 316

RESULT 10
 PICE_HUMAN STANDARD: PRT: 364 AA.
 AC Q9N0Q2; Q81Z47; Q9B0G4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 1-acyl-sn-glycerol 3-phosphate acyltransferase epsilon (BC 2.3.1.51)
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid
 DE acyltransferase-epsilon) (LPAT-epsilon) (1-acylglycerol-3-phosphate
 DE O-acyltransferase 5).
 GN AGPAT5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Leung D.W.;
 RT "Cloning and expression of LPAT-epsilon."
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;
 RL Patent number WO9932644, 01-JUL-1999.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Amygdala;
 RX Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glasel S.,
 RA Ansoerg W., Boecker M., Bloeker H., Bauersachs S., Blum H.,

RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Oltmannseder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.,
 RA "Towards a catalog of human genes and proteins: sequencing and
 RA analysis of 500 novel complete protein coding human cDNAs.",
 RA Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuma M., Hosolri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Niimiya K., Iwayanagi T.,
 RA "NEO human cDNA sequencing project.",
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RA and mouse cDNA sequences.",
 RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
 CC -1- PATHWAY: De novo phospholipid biosynthesis; second step.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
 CC ACYLTRANSFERASE FAMILY.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-12 is the initiator.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF375789; AK54809.1; -
 CC EMBL: AL136587; CAB6522.1; -
 CC EMBL: AK002072; BAA92069.1; ALT_INIT.
 CC EMBL: BC023550; AAH23550.1; -
 CC InterPro: IPR002123; Acyltransferase.
 CC Pfam: PF01553; Acyltransferase; 1.
 CC SMART: SM00563; Plac. 1.
 CC Phospholipid biosynthesis; Transferase; Acyltransferase;
 CC Transmembrane.
 KM TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 61 81 POTENTIAL.
 FT TRANSMEM 344 364 POTENTIAL.

FT CONFLICT 156 156 L -> V (IN REF. 2 AND 4).
 SQ SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C78081 CRC64;
 Query Match 45.9%; Score 39; DB 1; Length 364;
 Best Local Similarity 35.7%; Pred. No. 22;
 Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KSSQSVLYSANRKN 14
 Db 82 KKNENITVLANHQS 95
 : : : : :
 ID SYG_BOMMO STANDARD; PRT; 680 AA.
 AC 004451;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine-tRNA ligase) (GLYRS).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OC NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93216727; PubMed=8463296;
 RA Nada S., Chang P.K., Dignam J.D.;
 RT "Primary structure of the gene for glycyl-tRNA synthetase from Bombyx
 RT mori".
 RT J. Biol. Chem. 268:7660-7667(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate
 CC + glycyl-tRNA(Gly).
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC -1- SIMILARITY: Contains 1 WHEP-TRS domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: I08106; AA62231.1; -
 CC HSSP: P56206; IAT1.
 CC InterPro: IPR004154; HGRP_antlicodon.
 CC InterPro: IPR002314; tRNA-synt_2b.
 CC InterPro: IPR002315; tRNA-synt_gly.
 CC InterPro: IPR006195; tRNA_ligase_II.
 CC InterPro: IPR000738; WHEP-TRS.
 CC Pfam: PF03129; HGRP_antlicodon; 1.
 CC Pfam: PF00587; tRNA-synt_2b; 1.
 CC Pfam: PF00458; WHEP-TRS; 1.
 CC PRINTS: PR01043; TRNASYNTGLY.
 CC TIGRPFAM: TIGR00389; glys_dimeric; 1.
 CC PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
 CC PROSITE: PS00762; WHEP-TRS; 1.
 CC KMW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 20 65 WHEP-TRS.
 SQ SEQUENCE 680 AA; 76920 MW; 59118DA04922672F CRC64;
 QY 1 KSSQSVLYSANRKN 13
 Db 313 KNTQSLVYASADNQ 325
 : : : : :
 Query Match 45.9%; Score 39; DB 1; Length 680;
 Best Local Similarity 53.8%; Pred. No. 43;
 Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
 SM4F_HUMAN STANDARD; PRT; 770 AA.
 ID 095754; Q9NS35;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4F precursor (Semaphorin W) (Sema W) (Semaphorin M) (Sema M).
 GN SEMA4F OR SEMA4F OR SEMA4F.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM SMALL).
 RP TISSUE-Brain.
 RX MEDLINE=9916263; PubMed=10051670;
 RA Ekins J.A., Kikuchi K., Chedotal A., de Castro F., Goodman C.S., Kimura T.;
 RT "Cloning, expression, and genetic mapping of Sema W, a member of the semaphorin family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2491-2496(1999).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Chen F., Do A., Do T., Weisler M., Roe B.A.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE-Amygdala;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Giesel S., Ansgor W., Boecker M., Bloeker H., Bauersachs S., Blum H., Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N., Mewes H.-W., Ottenwelder B., Oberwiler B., Tampe J., Heubner D., Wambut R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 [4]
 RP SEQUENCE OF 347-770 FROM N.A.
 RA Jang W., Spilson S.V., Hua A., Roe B., Weisler M.H.;
 RT "Large-scale comparative sequence analysis of human and mouse genomic DNA in the mm2 region of mouse chromosome 6 reveals coding regions of three new genes.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: HAS GROWTH CONE COLLAPSE ACTIVITY AGAINST RETINAL GANGLION-CELL AXONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms-2;
 CC Name-Long;
 CC IsoId=095754-1; Sequence-Displayed;
 CC Name-Short;
 CC IsoId=095754-2; Sequence-VSP_006043;
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL; AB022317; BAA75631.1; -
 DR EMBL; AC006544; -; NOT_ANNOTATED_CDS.
 DR EMBL; AL136552; CAB66487.1; -
 DR EMBL; AF053369; AAF80660.1; -
 DR Gene; HGNC:10734; SEMA4F.
 DR MIM; 603706; -

DR GO: 0005887; C: integral to plasma membrane; TAS.
 DR GO: 0005624; C: membrane fraction; TAS.
 DR GO: 0007411; P: axon guidance; TAS.
 DR GO: 0007267; P: cell-cell signaling; TAS.
 DR GO: 0007399; P: neurogenesis; TAS.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01437; PSI; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein;
 KW Alternative splicing.
 FT SIGNAL 1 34 POTENTIAL.
 FT CHAIN 35 770 SEMAPHORIN 4F.
 FT DOMAIN 35 659 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 660 680 POTENTIAL.
 FT DOMAIN 681 770 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 65 496 SEMA.
 FT DOMAIN 512 563 PSI.
 FT DOMAIN 580 635 IG-LIKE C2-TYPE.
 FT DISULFD 587 628 BY SIMILARITY.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 509 509 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPIC 120 274 Missing (in isoform short).
 FT FTID=VSP_006043.
 FT CONFLICT 533 533 S -> N (IN REF. 1).
 SQ SEQUENCE 770 AA; 83511 MW; CFB974B41DF09C8 CRC64;
 Query Match 45.9%; Score 39; DB 1; Length 770;
 Best Local Similarity 72.7%; Pred. No. 50;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 6 VLYSANRKNRYL 16
 DB 186 VLYAATVKNRYL 196
 RESULT 13
 CDH_PHACH STANDARD; PRT; 773 AA.
 ID 001738; O00047;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cellulose dehydrogenase precursor (EC 1.1.99.18) (CDH) (Cellulose-DE quinone oxidoreductase).
 GN CDH-1 AND CDH-2.
 OS Phanerochaete chrysosporium.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Ascomycota; Corticiaceae; Phanerochaete.
 NC NCBI_TaxID=5306;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OGC101;
 RX MEDLINE=97077226; PubMed=8919793;
 RA Li B., Nagalla S.R., Renganathan V.;
 RT "Cloning of a cDNA encoding cellulose dehydrogenase, a hemoflavoenzyme from Phanerochaete chrysosporium.";
 RL Appl. Environ. Microbiol. 62:1329-1335(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OGC101;
 RX MEDLINE=97176414; PubMed=9023960;
 RA Li B., Nagalla S.R., Renganathan V.;
 RT "Cellulose dehydrogenase from Phanerochaete chrysosporium is encoded by two allelic variants.";
 RL Appl. Environ. Microbiol. 63:796-799(1997).
 [3]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.

RX MEDLINE-20139694; PubMed-10673428;
 RA Halberg B.M., Bergfors T., Boeckro K., Petersson G., Henriksson G.,
 RA Digne C.;
 RT "A new scaffold for binding haem in the cytochrome domain of the
 RT extracellular flavocytochrome cellobiose dehydrogenase.";
 RL Structure 8:79-88(2000).
 CC -1- FUNCTION: Degrades both lignin and cellobiose. Oxidizes cellobiose
 CC to cellobionolactone.
 CC -1- CATALYTIC ACTIVITY: cellobiose + acceptor = cellobiono-1,5-lactone
 CC + reduced acceptor.
 CC -1- COFACTOR: Binds 1 FAD, 1 HEME and 1 Iron ion per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE GMC
 CC OXIDOREDUCTASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U46081; AAC49277.1; -
 CC EMBL: U65888; AAB61455.1; -
 CC EMBL: U50409; AAB92262.1; -
 CC PDB: 1D7B; 18-OCT-99.
 CC PDB: 1D7C; 18-OCT-99.
 CC PDB: 1D7D; 18-OCT-99.
 CC PDB: 1KDG; 13-NOV-02.
 CC PDB: 1NAA; 14-JAN-03.
 CC InterPro: IPR000172; GMC_oxred.
 CC InterPro: IPR001100; Pyr_redox.
 CC Pfam: PF05199; GMC_oxred_C; 1.
 CC Pfam: PF00732; GMC_oxred_N; 1.
 CC PRINTS: PR00411; PNDRTASE1.
 CC PROSITE: PS00623; GMC_OXRED_1; 1.
 CC PROSITE: PS00624; GMC_OXRED_2; 1.
 CC KMO Cellulose degradation: Oxidoreductase; FAD; Flavoprotein; Heme;
 CC Multigene family; Signal; 3D-structure.
 CC STGNL 1 18
 CC FT CHAIN 19 773 CELLOBIOSE DEHYDROGENASE.
 CC FT DOMAIN 19 208 HEME DOMAIN.
 CC FT METAL 83 83 OXIDOREDUCTASE.
 CC FT METAL 181 181 IRON (HEME AXIAL LIGAND).
 CC FT NE_BIND 236 265 IRON (HEME AXIAL LIGAND).
 CC FT STRAND 20 20 FAD (ADP PART) (POTENTIAL).
 CC FT STRAND 22 25
 CC FT TURN 27 29
 CC FT STRAND 32 38
 CC FT STRAND 39 42
 CC FT STRAND 43 49
 CC FT STRAND 62 69
 CC FT TURN 70 71
 CC FT STRAND 74 78
 CC FT STRAND 79 80
 CC FT STRAND 88 94
 CC FT TURN 95 96
 CC FT STRAND 97 104
 CC FT STRAND 112 113
 CC FT STRAND 119 122
 CC FT TURN 124 125
 CC FT STRAND 127 128
 CC FT STRAND 132 140
 CC FT TURN 141 141
 CC FT STRAND 144 144
 CC FT TURN 146 147
 CC FT STRAND 150 150
 CC FT TURN 153 154
 CC FT STRAND 156 164
 CC FT TURN 171 172
 CC FT TURN 174 175

FT STRAND 183 190
 FT HELIX 191 193
 FT STRAND 195 195
 FT TURN 197 198
 FT HELIX 199 203
 FT TURN 204 205
 SQ SEQUENCE 773 AA; 82007 MW; 54F721E779AA4D7B CRC64;
 Query Match 45.9%; Score 39; DB 1; Length 773;
 Best Local Similarity 58.3%; Pred. No. 50;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 5 SVLYSNNHNYL 16
 Db 192 STAHSAANYNYL 203
 ID KYZE_HUMAN STANDARD; PRT; 117 AA.
 AC P06309;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-II region G607 precursor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_Taxid-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84191506; PubMed-6325927;
 RA Klobbeck H.G., Solomon A., Zachau H.G.;
 RT "Contribution of human V kappa II germ-line genes to light-chain
 RT diversity.";
 RL Nature 309:73-76(1984).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z00009; -; NOT_ANNOTATED_CDS.
 CC DR PIR: A01889; K2HDCM.
 CC DR HSSP: P80362; IWTL.
 CC DR GO: GO:0005576; C:extracellular; NAS.
 CC DR GO: GO:0003823; F:antigen binding activity; NAS.
 CC DR GO: GO:0006955; P:immune response; NAS.
 CC DR InterPro: IPR007110; Ig_Like.
 CC DR InterPro: IPR003006; Ig_MHC.
 CC DR Pfam: PF00047; Ig_V.
 CC DR SMART: SM00406; IGV_1.
 CC DR PROSITE: PS0835; IG_Like; 1.
 CC KW Immunoglobulin V region; Signal.
 CC FT NON_TER 1 1
 CC FT SIGNAL 1 4
 CC FT CHAIN 5 117 IG KAPPA CHAIN V-II REGION G607.
 CC FT DOMAIN 5 27 COMPLEMENTARITY-DETERMINING-1.
 CC FT DOMAIN 28 43 FRAMEWORK-2.
 CC FT DOMAIN 44 58 COMPLEMENTARITY-DETERMINING-2.
 CC FT DOMAIN 59 65 FRAMEWORK-3.
 CC FT DOMAIN 66 97 COMPLEMENTARITY-DETERMINING-3.
 CC FT DOMAIN 98 106 FRAMEWORK-4.
 CC FT DOMAIN 107 116 BY SIMILARITY.
 CC FT DISULFID 27 97
 CC FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 12664 MW; 92C57DC719E558B1 CRC64;
 Query Match 45.3%; Score 38.5; DB 1; Length 117;

Best Local Similarity 62.5%; Pred. No. 7.5;
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 KSSQSVLYSANHKNYL 16
:||||:| | | | |
Db 28 RSSQSLHLS-NGNYL 42

RESULT 15

RECX_CLOAB

ID RECX_CLOAB STANDARD; PRT; 214 AA.

AC 097GF7;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

GN RECX OR CAC2410.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC SRRATN-ATCC 824 / DSM 792 / VKM B-1787;

RA MEDLINE-21359325; PubMed-11466286;

RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hill J., Wolf Y.I.,

RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum.

RL J. Bacteriol. 183:4823-4838(2001).

CC - FUNCTION: Modulates recA activity (By similarity).

CC - SUBCELLULAR LOCATION: Cytoplasmic (Potential).

CC - SIMILARITY: BELONGS TO THE RECX FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE007741; AAK80365.1; -

DR PIR; B97197; B97197.

DR HAMAP; MF_01114; -; 1.

DR InterPro: IPR003783; RecX.

DR Pfam; PF02631; RecX; 1.

KW Complete proteome.

SQ SEQUENCE 214 AA; 25501 MW; 3957D81B7FE9B4AB CRC64;

QY 8 YSANHKNYL 16
:||||:| | | | |
Db 152 YSANNEYI 160

Query Match 44.7%; Score 38; DB 1; Length 214;

Best Local Similarity 66.7%; Pred. No. 18;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Search completed: October 7, 2003, 19:14:57
Job time : 5.45862 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 16.9414 Seconds

(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-2_COPY_24_40

Perfect score: 85

Sequence: 1 KSSQSVLYSANHKNTYIA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	56	65.9	107	11	Q9ERZ9	Q9ERZ9 mus musculi
2	50	58.8	130	4	Q9NP29	Q9NP29 homo sapien
3	45	52.9	146	8	Q8MOE1	Q8MOE1 cryptococcu
4	44	51.8	378	16	Q8UIX2	Q8UIX2 agrobacteri
5	43	50.6	306	5	Q9BK13	Q9BK13 paramecium
6	43	50.6	1360	4	Q9HCK2	Q9HCK2 homo sapien
7	43	50.6	1746	4	Q8NHE1	Q8NHE1 homo sapien
8	43	50.6	1833	4	Q9NR92	Q9NR92 homo sapien
9	43	50.6	2012	5	Q8IDC5	Q8IDC5 plasmodium
10	43	50.6	2316	4	Q8WXA6	Q8WXA6 homo sapien
11	43	50.6	2342	4	Q8NG31	Q8NG31 homo sapien
12	42	49.4	90	12	Q4Z047	Q4Z047 human respi
13	42	49.4	160	9	Q8SDC2	Q8SDC2 pseudomonas
14	42	49.4	291	5	Q94467	Q94467 dictyostell
15	42	49.4	291	16	Q8XNP9	Q8XNP9 clostridial
16	42	49.4	330	16	Q25502	Q25502 helicobacte

17	42	49.4	415	10	Q9SCW7	Q9SJW7 arabidopsis
18	42	49.4	708	5	Q9W0D8	Q9W0D8 drosophila
19	42	49.4	779	10	Q9LNU0	Q9LNU0 arabidopsis
20	42	49.4	2672	5	Q8IEE4	Q8IEE4 plasmodium
21	41	48.2	207	2	Q46674	Q46674 escherichia
22	41	48.2	444	11	Q9D2Q3	Q9D2Q3 mus musculi
23	41	48.2	480	5	Q8IND4	Q8IND4 drosophila
24	41	48.2	518	5	Q8I3Q2	Q8I3Q2 plasmodium
25	41	48.2	1351	5	Q8T6H3	Q8T6H3 dictyostell
26	40.5	47.6	119	2	Q9EV97	Q9EV97 streptococc
27	40.5	47.6	297	16	Q8R6A4	Q8R6A4 fusobacteri
28	40	47.1	365	11	Q8CGN6	Q8CGN6 mus musculi
29	40	47.1	365	11	Q8BG61	Q8BG61 mus musculi
30	40	47.1	406	16	Q8K8Q8	Q8K8Q8 streptococc
31	40	47.1	410	16	Q9A1L2	Q9A1L2 streptococc
32	40	47.1	410	16	Q8P2R6	Q8P2R6 streptococc
33	40	47.1	425	5	Q8MYE8	Q8MYE8 dictyostell
34	40	47.1	469	16	Q8NPY6	Q8NPY6 corynebacte
35	40	47.1	488	17	Q979J5	Q979J5 thermoplasma
36	40	47.1	506	10	Q4O140	Q4O140 lycopersico
37	40	47.1	510	8	Q92Y12	Q92Y12 papilio ind
38	40	47.1	542	4	Q9NSG4	Q9NSG4 homo sapien
39	40	47.1	543	5	Q8I345	Q8I345 plasmodium
40	40	47.1	564	5	Q8IS03	Q8IS03 drosophila
41	40	47.1	673	4	Q9NSG3	Q9NSG3 homo sapien
42	40	47.1	674	4	Q9NWZ7	Q9NWZ7 homo sapien
43	40	47.1	674	4	Q9BTX1	Q9BTX1 homo sapien
44	40	47.1	702	5	Q95PI7	Q95PI7 tenebrio mo
45	40	47.1	797	10	Q9SK37	Q9SK37 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q9ERZ9	PRELIMINARY:	PRT:	107 AA.
AC	Q9ERZ9	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	Anti human TNF-alpha light chain variable region (Fragment).				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
PP	SEQUENCE FROM N.A.				
RA	Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;				
RT	"Cloning and sequencing of the light chain fragment of variable region				
RT	genes of an anti-hTNF-a monoclonal antibody.";				
RL	J. Cell. Mol. Immunol. 12:21-26(1996).				
RN	[2]				
PP	SEQUENCE FROM N.A.				
RA	Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;				
RT	"Construction and sequencing of the single-chain antibody gene of a				
RT	human TNF-alpha specific monoclonal antibody.";				
RL	Ti 4 Chun i Ta Hsueh Hsueh Pao 19:373-376(1998).				
RN	[3]				
PP	SEQUENCE FROM N.A.				
RA	Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;				
RT	Submitted (May-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF262753; AAG23804.1; -				
DR	HSSP; P80362; IWTI.				
DR	InterPro: IPR007110; Ig-like.				
DR	InterPro: IPR003006; Ig_MHC.				
DR	InterPro: IPR003596; Ig_v.				
DR	Pfam: PF00047; Ig; 1.				
DR	SMART; SM00406; IGV; 1.				
DR	PROSITE; PS50835; IG_LIKE; 1.				
FT	NON_TER				
FT	SEQUENCE				
FT	107 AA; 11784 MW; 2B15EBA6604A26C3 CRC64;				

Query Match 65.9%; Score 56; DB 11; Length 107;
 Best Local Similarity 76.5%; Pred. No. 0.1;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 KSSOSVLSANHKNYLA 17
 ||||| | |||||
 DB 21 KSSOSVLSNTRKNYLA 37

RESULT 2

O9NP29 PRELIMINARY; PRT; 130 AA.
 AC O9NP29;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Microfibrillar protein 2 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=96354815; PubMed=8753791;
 RA Ozavath K.J., Xia S., Hirose H., Tilson M.D.;
 RT "Two hypothetical proteins of human aortic adventitia, with Ig kappa,
 collagenous, and aromatic-rich motifs".
 RL Biochem. Biophys. Res. Commun. 225:500-504(1996).
 RN [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97367690; PubMed=9224393;
 RA Ozavath K.J., Hirose H., Xia S., Chew D., Knoetgen J. III,
 RA Tilson M.D.;
 RT "Expression of two novel recombinant proteins from aortic adventitia
 (kappaIb) sharing amino acid sequences with cytomegalovirus.";
 RL J. Surg. Res. 69:277-282(1997).
 RN [3]
 RP SEQUENCE FROM N.A.

RX Ozevath K.J., Xia S., Hirose H., Tilson M.D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF206020; AAF62402.1; -
 DR HSSP; P80362; 1MTL.

FT NON_TER 130
 SQ SEQUENCE 130 AA; 14128 MW; 51275185AAC6FA1E CRC64;

Query Match 58.8%; Score 50; DB 4; Length 130;
 Best Local Similarity 52.9%; Pred. No. 1.3;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KSSOSVLSANHKNYLA 17
 :||| : : : |||||
 DB 44 RSSQRLFGSNSKNYLA 60

RESULT 3

O8MOE1 PRELIMINARY; PRT; 146 AA.
 AC O8MOE1;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NADH dehydrogenase subunit 2 (Fragment).
 GN NAD2.
 OS Cryptococcus neoformans var. neoformans.
 OG Mitochondrion.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=40410;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-IFM 5844;
 RA Litterer J., Kucsera J.;
 RT "Partial sequence of the NADH dehydrogenase subunit 2 gene and the

RT NADH dehydrogenase subunit 3 gene of Cryptococcus neoformans var.
 neoformans IFM 5844.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF534567; AAN01219.1; -
 DR InterPro; IPR001750; Oxidored_q1.
 DR Pfam; PF00361; Oxidored_q1.1.
 KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
 FT NON_TER 1

SQ SEQUENCE 146 AA; 15835 MW; 83782F70D64DC4E8 CRC64;

Query Match 52.9%; Score 45; DB 8; Length 146;
 Best Local Similarity 57.1%; Pred. No. 9.6;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 SSOSVLSANHKNY 15
 :||| | |
 DB 52 AKQVLSATHSYG 65

RESULT 4

O8UIX2 PRELIMINARY; PRT; 378 AA.
 AC O8UIX2;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein phm.
 GN PHM OR ATU0169 OR AGR_C_279.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Gurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houtmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy D., Scott C., Lapps C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cleio C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008991; AAL41193.1; -
 DR EMBL; AE007958; AAK65988.1; -
 DR InterPro; IPR006680; Amidohydro_1.
 DR Pfam; PF01979; Amidohydro_1.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 378 AA; 41200 MW; 81EB8B837A75EB45 CRC64;

Query Match 51.8%; Score 44; DB 16; Length 378;
 Best Local Similarity 47.1%; Pred. No. 37;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 KSSOSVLSANHKNYLA 17
 : : | ||| | : |||||

DB 201 RQASQVLYSANHNYIA 217

RESULT 5
09BK13
ID 09BK13 PRELIMINARY; PRT; 306 AA.

AC 09BK13; 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN GAPC.
OS Paramecium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
OC Paramecium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21143069; PubMed-11230543;
RA Fast N.M., Kissinger J.C., Roos D.S., Keeling P.J.;
RT "Nuclear-Encoded, Plastid-Targeted Genes Suggest a Single Common
Origin for Alveolates and Dinoflagellate Plastids."
RL Mol. Biol. Evol. 18:418-426(2001).
CC -1. SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DEHYDROGENASE FAMILY.
DR EMBL; AF319452; AKR20731.1; -.
DR HSSP; P56649; IDSS.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh_1.
DR PRINTS; PR00078; G3PDHGNASE.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT 306 306
SQ SEQUENCE 306 AA; 32657 MW; 680D7EC6EDA527C CRC64;

QY 1 KSSQSVLYSANHNYIA 17
DB 117 KGTPTFYGVNHNRYA 133

Query Match 50.6%; Score 43; DB 5; Length 306;
Best Local Similarity 41.2%; Pred. No. 44;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

RESULT 6
09HCK2
ID 09HCK2 PRELIMINARY; PRT; 1360 AA.

AC 09HCK2; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE Hypothetical protein KIAA1570 (Fragment).
GN KIAA1570.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20450683; PubMed-10997877;
RA Nagase T., Kikuno R., Nakayama M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
XVIII. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046790; BAB13396.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT 1360 1360
SQ SEQUENCE 1360 AA; 154971 MW; 7834D7E626D212B8 CRC64;

Query Match 50.6%; Score 43; DB 4; Length 1360;

Best Local Similarity 42.9%; Pred. No. 1.9e+02;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKN 14
DB 115 KNDKTIVFSENHKN 128

RESULT 7
08NH1
ID 08NH1 PRELIMINARY; PRT; 1746 AA.

AC 08NH1; 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, last annotation update)
DE D40 protein.
GN D40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-Testis;
RX MEDLINE-22081483; PubMed-12087463;
RA Takimoto M., Wei G., Dosaka-Akita H., Mao P., Kondo S., Sakuregi N.,
RA Chiba I., Miura T., Itoh N., Sasao T., Koya R.C., Tsukamoto T.,
RA Fujimoto S., Kato H., Kuzumaki N.;
RT "Frequent expression of new cancer/testis gene D40/AF15q14 in lung
cancer of smokers."
RL Br. J. Cancer 86:1757-1762(2002);
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-Testis;
RX MEDLINE-20241296; PubMed-10780384;
RA Wei G., Takimoto M., Yoshida I., Mao P., Koya R.C., Miura T.,
RA Kuzumaki N.;
RT "Chromosomal assignment of a novel huma gene D40."
RL Nucleic Acids Symp. Ser. 42:71-72(1999).
DR EMBL; AB022190; BAC05691.1; -.
SQ SEQUENCE 1746 AA; 195451 MW; 8183D802B8CE0991 CRC64;

Query Match 50.6%; Score 43; DB 4; Length 1746;
Best Local Similarity 42.9%; Pred. No. 2.5e+02;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKN 14
DB 1071 KNDKTIVFSENHKN 1084

RESULT 8
09NR92
ID 09NR92 PRELIMINARY; PRT; 1833 AA.

AC 09NR92; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE AF15q14 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20438193; PubMed-10980622;
RA Hayette S., Tigand I., Vanier A., Martel S., Corbo L., Charlin C.,
RA Bellard E., Deleage G., Magaud J.P., Rimokh R.;
RT "AF15q14, a novel partner gene fused to the MLL gene in an acute
myeloid leukaemia with a t(11;15)(q23;q14)."
RL Oncogene 19:4446-4450(2000).
DR EMBL; AF248041; AAF97513.1; -.
SQ SEQUENCE 1833 AA; 205683 MW; D13BBE8784B78369 CRC64;

Query Match 50.6%; Score 43; DB 4; Length 1833;
Best Local Similarity 42.9%; Pred. No. 2.6e+02;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKN 14
DB 1071 KNDKTIVSEHNKN 1084
|:|||||

RESULT 9
ID 08IDC5 PRELIMINARY; PRT: 2012 AA.
AC 08IDC5.
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN M13p1.285.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Bartell B.,
RL Submitted (Sep-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52700.1; -
KM Hypothetical protein.
SQ SEQUENCE 2012 AA; 238205 MW; 7A3C897267BAE718 CRC64;

Query Match 50.6%; Score 43; DB 5; Length 2012;
Best Local Similarity 35.3%; Pred. No. 2.9e+02;
Matches 6; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKN 17
DB 1092 KENRTVYLNKNKEYIS 1108
|:|||||

RESULT 10
ID 08WXA6 PRELIMINARY; PRT: 2316 AA.
AC 08WXA6.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE AF15q14 isoform 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20438193; PubMed=10980622;
RA Hayward S., Tliguid I., Vanter A., Martel S., Corbo L., Charrin C.,
RA Bellard E., Deleage G., Magaud J.P., Rimokh R.,
RT AF15q14, a novel partner gene fused to the ML gene in an acute
myeloid leukemia with a t(11;15)(q23;q14).";
RL Oncogene 19:4446-4450(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Hayette S., Rimokh R.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF461041; AAL67803.1; -
SQ SEQUENCE 2316 AA; 262517 MW; AD77046BFCB8321C CRC64;

Query Match 50.6%; Score 43; DB 4; Length 2316;
Best Local Similarity 42.9%; Pred. No. 3.3e+02;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKN 14

DB 1071 KNDKTIVSEHNKN 1084
|:|||||

RESULT 11
ID 08NG31 PRELIMINARY; PRT: 2342 AA.
AC 08NG31.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE AF15q14.
GN AF15Q14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuefer M.V., Chinwalla V., Zelezniuk-Le N.J., Behm F.G., Naeye C.W.,
RA Rakestraw K.M., Rainoldi S.C., Morris S.W.,
RT "Characterization of the ML partner gene, AF15q14, involved in the
t(11;15)(q23;q14-15).";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF173994; AAM45143.1; -
SQ SEQUENCE 2342 AA; 265284 MW; 6844183244CCB86B CRC64;

Query Match 50.6%; Score 43; DB 4; Length 2342;
Best Local Similarity 42.9%; Pred. No. 3.4e+02;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSSOSVLYSANHKN 14
DB 1097 KNDKTIVSEHNKN 1110
|:|||||

RESULT 12
ID 042047 PRELIMINARY; PRT: 90 AA.
AC 042047.
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Matrix protein 2.
GN M2.
OS Human respiratory syncytial virus (strain B1).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=79692;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B1;
RA Karron R.A., Buongiorno D.A., Georgiu A.F., Whitehead S.S.,
RA Adamus J.E., Clements-Mann M.L., Harris D.O., Randolph V.B.,
RA Uden S.A., Murphy B.R., Sidhu M.S.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-B1;
RA Sidhu M.S., Uden S.A., Murphy B.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF013254; AAB82438.1; -
DR EMBL; AF013255; AAB82448.1; -
KM Matrix protein.
SQ SEQUENCE 90 AA; 10587 MW; 16D886E983DE3994 CRC64;

Query Match 49.4%; Score 42; DB 12; Length 90;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSOSVLYSANHKN 16
DB 23 SSESMTATFNHKN 37
|:|||||

RESULT 13

08SDC2
ID 08SDC2 PRELIMINARY; PRT; 160 AA.
AC 08SDC2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE PHIK2040.
OS Pseudomonas phage phiK2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=169683;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21914557; PubMed=11916376;
RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostynchenko V.A.,
Bouratalsseva M.V., Syktilinda N.N., Kiylov V.V., Volckaert G.,
"The genome of bacteriophage phiK2 of Pseudomonas aeruginosa.",
J. Mol. Biol. 317:1-19(2002).
RL [2]
RN SEQUENCE FROM N.A.
RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostynchenko V.A.,
Bouratalsseva M.V., Syktilinda N.N., Kiylov V.V., Volckaert G.,
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: A399011; AAL82941.1; -
DR InterPro: IPR000387; TYR_phosphatase.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 160 AA; 18179 MW; 189BDD24A3992106 CRC64;
Query Match 49.4%; Score 42; DB 9; Length 160;
Best Local Similarity 80.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 7 LYSANHKNTL 16
DB 35 LYSAKHLNLYL 44
RESULT 14
094467
ID 094467 PRELIMINARY; PRT; 291 AA.
AC 094467;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE SRNA-binding protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RX MEDLINE=98086164; PubMed=9426195;
RA Oberosler P., Nellen W.;
RT "Functional activity and developmental regulation of DdRBP1, a RNA
binding protein in Dictyostelium discoideum.",
Mol. Biol. Chem. 378:1353-1360(1997).
DR EMBL: Y07952; CAA69250.1; -
DR HSSP: P09651; IHAL.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 291 AA; 31767 MW; D8B717C82AE0C2D5 CRC64;
Query Match 49.4%; Score 42; DB 5; Length 291;
Best Local Similarity 62.5%; Pred. No. 61;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSQSIVLYSANHKNTL 17
DB 125 SSRKVTYSANTINMLA 140

RESULT 15

08XNP9
ID 08XNP9 PRELIMINARY; PRT; 291 AA.
AC 08XNP9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Hypothetical protein CPE0283.
GN CPE0283.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
"Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater.",
Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003186; BAB79989.1; -
DR InterPro: IPR001823; Aldol-epimerase.
DR Pfam: PF01263; Aldose_epim; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 291 AA; 34015 MW; E2B6D485064CFE38 CRC64;
Query Match 49.4%; Score 42; DB 16; Length 291;
Best Local Similarity 56.2%; Pred. No. 61;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KSSQSVLYSANHKNTL 16
DB 211 KSNKISLKSNNHKNTL 226

Search completed: October 7, 2003, 19:19:55
Job time : 18.9414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:26 ; Search time 8.81034 Seconds
(without alignments)
126.112 Million cell updates/sec

Title: US-09-988-013a-2_COPY_56_62
Perfect score: 38
Sequence: 1 WASTRES 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_19Jun03: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	14 AAR30157	MAB GAH variable r
2	38	100.0	7	15 AAR57964	MAB NFE2 light cha
3	38	100.0	7	15 AAR57966	MAB NFE2 light cha
4	38	100.0	7	15 AAR50317	MAB NFE2 light cha
5	38	100.0	7	20 AAY08713	Human milk fat glo
6	38	100.0	7	21 AAB07952	CDR2 sequence from
7	38	100.0	7	21 AAY95234	Human monoclonal a
8	38	100.0	7	21 AAY32227	G-CSF agonist anti
9	38	100.0	7	23 AAE29275	Anti-ClfA monoclon

10	38	100.0	7	23 ABE62361	Human immunopeptid
11	38	100.0	7	23 ABE68854	CDR2 VL peptide.
12	38	100.0	7	23 AAB81257	Human trkC antibod
13	38	100.0	7	23 AAU70335	Human Kappa IV lig
14	38	100.0	7	23 AAU70347	Mouse Kappa I lig
15	38	100.0	7	24 ABE98681	Human GAH antibody
16	38	100.0	7	24 ABE98689	Human GAH antibody
17	38	100.0	7	24 ABE34369	Escherichia coli I
18	38	100.0	7	24 ABE34369	CDR2 of HBEG-1 lig
19	38	100.0	15	23 ABE68859	CDR2 sequence remo
20	38	100.0	82	17 AAE62807	Amino acid sequenc
21	38	100.0	84	17 AAE14491	Monoclonal antibod
22	38	100.0	84	17 AAE98978	Monoclonal antibod
23	38	100.0	94	21 AAE56699	Partial peptide fr
24	38	100.0	101	24 ABE18696	Antibody library r
25	38	100.0	105	23 ABE07184	MS11955 light cha
26	38	100.0	107	19 AAE80817	Amino acid sequenc
27	38	100.0	107	19 AAE80814	Amino acid sequenc
28	38	100.0	107	20 AAE17955	Human kappa 8 lig
29	38	100.0	110	13 AAE30012	Light chain variab
30	38	100.0	110	19 AAE56520	Anti-DNA IL/IM (R2
31	38	100.0	110	19 AAE56517	H221 antibody VL r
32	38	100.0	111	15 AAE52059	Light chain variab
33	38	100.0	112	13 AAE28807	5A8 VL. Synthetic
34	38	100.0	112	15 AAE54932	Fc receptor humant
35	38	100.0	112	15 AAE54933	MAB 022 VR chain.
36	38	100.0	112	23 AAE29264	S. aureus ClfA spe
37	38	100.0	112	23 AAE29266	S. aureus ClfA spe
38	38	100.0	112	23 AAE29268	S. aureus ClfA spe
39	38	100.0	112	23 AAE29270	S. aureus ClfA spe
40	38	100.0	112	23 ABE76514	HCV EI antigen mon
41	38	100.0	112	23 ABE76514	HCV EI antigen mon
42	38	100.0	112	23 ABE76548	HCV EI antigen mon
43	38	100.0	113	13 AAE25407	Light chain variab
44	38	100.0	113	15 AAE50313	Humanised light ch
45	38	100.0	113	15 AAE50314	Humanised light ch

ALIGNMENTS

RESULT 1	
AA30157	AA30157 standard; Protein: 7 AA.
XX	XX
AC	AA30157:
XX	XX
DT	25-MAR-2003 (updated)
DT	06-MAY-1993 (first entry)
XX	XX
DE	MAB GAH variable region of light chain.
XX	XX
KW	Monoclonal antibody; hybridoma; PCR; variable region;
KW	constant region; heavy chain; light chain.
XX	XX
OS	Synthetic.
PN	EP520499-A1.
XX	XX
PD	30-DEC-1992.
XX	XX
PF	26-JUN-1992; 92EP-0110841.
XX	XX
PR	28-JUN-1991; 91JP-0158859.
PR	28-JUN-1991; 91JP-0158860.
PR	28-JUN-1991; 91JP-0158861.
XX	XX
PA	(MITU) MITSUBISHI KASEI CORP.
XX	XX
PI	Hirakawa Y, Hosokawa S, Ito N, Nagaike K, Tagawa T;
XX	XX
DR	WPI; 1993-001328/01.
DR	N-PSDB; AAO33044.

XX Human monoclonal antibody specific for a cancer cell membrane
PT surface antigen - prep. from a hybridoma obtd. by cell fusion
PT between human lymphocytes derived from cancer patients and mouse
PT myeloma cells
PS
XX
PS Claim 3; Page 30 + 23; 37pp; English.
XX
XX The sequence is described is the specification as having 27 bases.
CC A human Mab specifically binding to a surface antigen of cancer
CC cell membrane comprises variable regions of the heavy and light
CC chains having the amino acid sequences of AAR30153-55 and AAR30156-58
CC respectively, encoded by DNA sequences AA033040-42 and AA033043-45
CC respectively. The antibody is obtained from a hybridoma producing
CC human antibody GAN.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 38; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WASTRES 7
DB 1 WASTRES 7
DB
RESULT 2
AAR57964
ID AAR57964 standard; Protein: 7 AA.
XX
XX AAR57964;
XX
XX 25-MAR-2003 (updated)
DT 05-OCT-1994 (first entry)
XX
XX MAB NFS2 light chain variable region CDR2.
DE
XX
XX Monoclonal antibody; Plasmidium falciparum; CDR;
KW complementarity determining region; fusion protein;
KW murine; variable; light; heavy; chain; malaria.
XX
XX Plasmidium falciparum.
OS
XX
XX WO9405690-A1.
PN
XX
XX 17-MAR-1994.
PD
XX
XX 08-SEP-1993; 93WO-US08435;
PF
XX
XX 09-SEP-1992; 92US-0941654.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (USNA) US SEC OF ARMY.
PA (USNA) US SEC OF NAVY.
XX
XX Charoenvit Y, Hoffman S, Hurle M, Rosenberg M;
PI Sadoff JC, Sylvester DR, Gross MS;
XX
XX MPI: 1994-101115/12.
DR N-PSDB: AA044837.
XX
XX
XX New engineered antibodies and fusion proteins for preventing
PT Plasmidium infection - contg. murine antibody CDR sequences, and
PT corresp. nucleic acid, vectors and transformed cells
XX
XX
PS Disclosure: Page 68; 98pp; English.
XX
XX Naturally-occurring (AA044841-42) and synthetic (AA044825-28)
CC variable light chain and variable heavy chain sequences
CC derived from murine mab NFS2 are provided. Murine mab NFS2,
CC its variable chain peptides, CDRs, functional fragments,
CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmidium infections.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 38; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WASTRES 7
DB 1 WASTRES 7
DB

CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmidium infections.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 38; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WASTRES 7
DB 1 WASTRES 7
DB
RESULT 3
AAR57966
ID AAR57966 standard; Protein: 7 AA.
XX
XX AAR57966;
XX
XX 25-MAR-2003 (updated)
DT 05-OCT-1994 (first entry)
XX
XX MAB NFS2 light chain variable region CDR2 (modified).
DE
XX
XX Monoclonal antibody; Plasmidium falciparum; CDR;
KW complementarity determining region; fusion protein;
KW murine; variable; light; heavy; chain; malaria.
XX
XX Synthetic.
OS
XX
XX WO9405690-A1.
PN
XX
XX 17-MAR-1994.
PD
XX
XX 08-SEP-1993; 93WO-US08435.
PF
XX
XX 09-SEP-1992; 92US-0941654.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (USNA) US SEC OF ARMY.
PA (USNA) US SEC OF NAVY.
XX
XX Charoenvit Y, Hoffman S, Hurle M, Rosenberg M;
PI Sadoff JC, Sylvester DR, Gross MS;
XX
XX MPI: 1994-101115/12.
DR N-PSDB: AA044839.
XX
XX
XX New engineered antibodies and fusion proteins for preventing
PT Plasmidium infection - contg. murine antibody CDR sequences, and
PT corresp. nucleic acid, vectors and transformed cells
XX
XX
PS Disclosure: Page 69; 98pp; English.
XX
XX Naturally-occurring (AA044841-42) and synthetic (AA044825-28)
CC variable light chain and variable heavy chain sequences
CC derived from murine mab NFS2 are provided. Murine mab NFS2,
CC its variable chain peptides, CDRs, functional fragments,
CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmidium infections.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 7 AA;
SQ
Query Match 100.0%; Score 38; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WASTRES 7
DB 1 WASTRES 7
DB

```

Db          |||||
            1 WASTRES 7

RESULT 4
AAR50317
ID AAR50317 standard; Protein; 7 AA.
XX
AC AAR50317;
XX
DT 25-MAR-2003 (updated)
DT 05-OCT-1994 (first entry)
XX
DE MAb NFS2 light chain variable region CDR2 (modified).
XX
KM Monoclonal antibody; Plasmodium falciparum; CDR;
KM complementarity determining region; fusion protein;
KM murine; variable; light; heavy; chain; malaria.
XX
OS Synthetic.
XX
PN WO9405690-A1.
XX
PD 17-MAR-1994.
XX
PF 08-SEP-1993; 93WO-US08435.
XX
PR 09-SEP-1992; 92US-0941654.
XX
PA (SMR ) SMITHKLINE BECHAM CORP.
PA (USNA ) US SEC OF ARMY.
PA (USNA ) US SEC OF NAVY.
XX
PI Charenylt Y, Hoffman S, Hurle M, Rosenberg M;
PI Sedoff JC, Sylvester DR, Gross MS;
XX
DR WPI; 1994-10115/12.
DR N-PSDB; AAQ44833.
XX
PT New engineered antibodies and fusion proteins for preventing
PT Plasmodium infection - contg. murine antibody CDR sequences, and
PT corresp. nucleic acid, vectors and transformed cells
XX
PS Claim 16; Page 83; 98pp; English.
XX
CC Naturally-occurring (AAQ44841-42) and synthetic (AAQ44825-28)
CC variable light chain and variable heavy chain sequences
CC derived from murine mAb NFS2 are provided. Murine mAb NFS2,
CC its variable chain peptides, CDRs, functional fragments,
CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmodium infections.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 7 AA:

Query Match          100.0%; Score 38; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
   |||||
Db 1 WASTRES 7

RESULT 5
AAY08713
ID AAY08713 standard; protein; 7 AA.
XX
AC AAY08713;
XX
DT 10-AUG-1999 (first entry)
XX

```

```

DE Human milk fat globule antigen binding fragment 5.
XX
XX Immunoglobulin; immunospecific; variable domain; CDR; cancer antigen;
KM complementarity determining region; infectious disease agent antigen;
KM cellular receptor; infectious agent; chemotherapeutic agent; vaccine;
KM cancer; infectious disease; autoimmune disease; Rheumatoid arthritis;
KM ulcerative colitis; psoriasis; allergy; immunoassay reagent; diagnosis;
KM veterinary medicine; antibody; immunisation; antibody; immune tolerance;
KM human milk fat globule antigen; cryptic expression;
KM anti-idiotypic immune response.
XX
OS Homo sapiens.
XX
PN WO925378-A1.
XX
PD 27-MAY-1999.
XX
PF 13-NOV-1998; 98WO-US24302.
XX
PR 10-APR-1998; 98US-0081403.
PR 14-NOV-1997; 97US-0065716.
XX
PA (EURO-) EUROCELTIQUE SA.
XX
PI Burch RM;
XX
DR WPI; 1999-357555/30.
XX
PT Modified immunoglobulins including engineered binding site
XX
PS Claim 9; Page 76; 123pp; English.
XX
CC This invention describes a novel modified immunoglobulin (mig), or its
CC fragment, that binds immunospecifically to one member (M1) of a binding
CC pair comprises a variable domain (V) having at least one CDR of the
CC (complementarity determining region) that contains a portion of the
CC second member (M2) of the pair. This portion is not present in the
CC natural CDR and M1 is a cancer antigen, infectious disease agent antigen,
CC a cellular receptor for an infectious agent (provided the binding site is
CC not NAMP or NMDP) or a member of a receptor-ligand pair. mig (also
CC related molecules containing the same V region), optionally coupled to a
CC chemotherapeutic agent, are used in therapeutic and/or prophylactic
CC compositions, or vaccines, against cancers, infectious or autoimmune
CC diseases (Rheumatoid arthritis, ulcerative colitis or psoriasis) or
CC allergies. They can also be used as immunoassay reagents for diagnosing
CC these conditions. mig can be used in human or veterinary medicine. mig
CC have higher affinity or specificity than native antibodies. Manipulation
CC of CDR ensures binding specificity and avoids the unpredictable
CC immunization and screening procedures currently used. It is now possible
CC to generate antibodies against antigens that are inaccessible, by virtue
CC of immune tolerance or cryptic expression. mig neutralize antigens
CC directly or they induce an anti-idiotypic immune response. This sequence
CC represents a fragment of modified immunoglobulin that immunospecifically
CC binds to a binding pair in which the first member is the human milk fat
CC globule antigen.
XX
SQ Sequence 7 AA:

Query Match          100.0%; Score 38; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
   |||||
Db 1 WASTRES 7

RESULT 6
AAB07952
ID AAB07952 standard; peptide; 7 AA.
XX
AC AAB07952;
XX

```

```

DE 14-NOV-2000 (first entry)
XX CDR2 sequence from an antibody with affinity for B7 molecules.
XX
XX Complementarity determining region; CDR; antibody; B7 molecule; B7-1;
XX B7-2; humanised immunoglobulin; autoimmune disease; infectious disease;
XX inflammatory disorder; systemic lupus erythematosus; diabetes mellitus;
XX insulin; asthma; arthritis; inflammatory bowel disease; cancer;
XX inflammatory dermatitis; multiple sclerosis; transplant rejection;
XX proliferative disease; leukemia; lymphoma; anaemia; sickle-cell anaemia;
XX thalassemia; aplastic anaemia; myeloid dysplasia syndrome.
XX
OS Mus sp.
XX
XX WO200047625-A2.
XX
XX 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US03303.
XX
XX 12-FEB-1999; 99US-0249011.
XX 24-JUN-1999; 99US-0339596.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Co MS, Vasquez M, Carreno B, Gelinker AC, Collins M, Goldman S;
XX Gray GS, Knight A, O'hara D, Rup B, Veldman GM;
XX
XX WPI; 2000-524532/47.
XX
XX Humanised immunoglobulin having a binding specificity to B7-1 (derived
XX from ATCC PTA-263), or B7-2 (derived from ATCC CRL-1524) molecules,
XX modulates immune responses and can therefore treat e.g. autoimmune
XX diseases, infectious diseases -
XX
XX Disclosure: Page 28; 162pp; English.
XX
XX The present sequence represents a complementarity determining region
XX (CDR) 2 from the light chains of an murine antibody with having a
XX binding specificity to B7-2 molecules. The sequence is used to construct
XX humanised immunoglobulins, which comprise an antigen binding region of
XX non-human origin and a portion of a human immunoglobulin. The humanized
XX immunoglobulins are useful for treating autoimmune diseases, infectious
XX diseases, inflammatory disorders, systemic lupus erythematosus, diabetes
XX mellitus, insulin; asthma, arthritis, inflammatory bowel disease,
XX inflammatory dermatitis, and multiple sclerosis. The immunoglobulins are
XX also useful for treating a transplant recipient or preventing transplant
XX rejection in a transplant recipient, and treating proliferative disease
XX (leukemia, lymphoma and cancer), anaemia (sickle-cell anaemia,
XX thalassemia and aplastic anaemia), inborn errors of metabolism,
XX congenital immunodeficiency diseases, and myeloid dysplasia syndrome.
XX
XX Sequence 7 AA;
XX
XX Query Match 100.0%; Score 38; DB 21; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 WASTRES 7
XX 1 WASTRES 7
XX DB 1 WASTRES 7
XX
XX RESULT 7
XX AAY95234
XX ID AAY95234 standard; Peptide: 7 AA.
XX
XX AC AAY95234;
XX
XX DT 29-AUG-2000 (first entry)
XX
XX DE Human monoclonal antibody LEN VL CDR2.
XX

```

```

XX Human; monoclonal antibody; LEN; humanised antibody; CC49; HuCC49;
XX CDR; complementarity determining region; colon cancer;
XX tumor associated glycoprotein-72; TAG-72; tumour marker; carcinoma;
XX diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..6
XX FT /note= "specificity determining region"
XX
XX PN WO200026394-A1.
XX
XX PD 11-MAY-2000.
XX
XX PF 29-OCT-1999; 99WO-US25552.
XX
XX PR 31-OCT-1998; 98US-0106534.
XX 02-NOV-1998; 98US-0106757.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Kashmiri SVS, Padlan EA, Schlom J;
XX
XX WPI; 2000-365637/31.
XX
XX Chimeric variants of CC49 monoclonal antibodies useful for detecting
XX and treating cancers associated with the expression of the pancarcinoma
XX tumor-associated antigen TAG-72 -
XX
XX Disclosure: Fig 2; 76pp; English.
XX
XX The present sequence is that of complementarity determining region
XX 2 (L-CDR2) of the light chain variable region (VL) of human
XX monoclonal antibody LEN. The invention is directed toward
XX mouse-human chimeric variants of CC49 Mabs with minimal murine
XX content, to methods of making such variants, and to their
XX therapeutic application. Variants are provided of huCC49, a
XX humanised Mab formed by grafting hypervariable regions from murine
XX CC49 into VL and VH frameworks of human MAb5 LEN and 21/28' CL,
XX respectively, while retaining murine framework residues required
XX for integrity of the antigen combining site structure. HuCC49
XX binds to the human pancreaticoma tumor associated glycoprotein-72
XX (TAG-72), which is found on the surface of certain human tumors.
XX Novel variants of huCC49 of the invention have fewer than all 6
XX CDRs of CC49 present. Also provided are specifically determining
XX region (SDR) variants of huCC49 in which only SDRs of at least 1
XX CDR from CC49 are present. Particular variants of HuCC49 have
XX either L-CDR1 and/or L-CDR2 from human MAb LEN. These variants
XX have the same or 2-fold lower affinity constant than HuCC49. They
XX are used in claimed methods of treating cancer and for detecting
XX cancer cells that express TAG-72.
XX
XX Sequence 7 AA;
XX
XX Query Match 100.0%; Score 38; DB 21; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 WASTRES 7
XX 1 WASTRES 7
XX DB 1 WASTRES 7
XX
XX RESULT 8
XX AAY32227
XX ID AAY32227 standard; Peptide: 7 AA.
XX
XX AC AAY32227;
XX
XX DT 15-FEB-2000 (first entry)
XX
XX DE G-CSF agonist antibody MAb163-93 light chain variable region CDR2.
XX

```

```

XX Granulocyte colony stimulating factor receptor; G-CSF; mouse;
XX monoclonal antibody; agonist; screening; neutropenia; therapy;
XX complementarity determining region; CDR; Mab163-93.
XX
XX Mus musculus.
XX
XX MO9955735-A1.
XX
XX 04-NOV-1999.
XX
XX 30-APR-1999; 99MO-US09466.
XX
XX 30-APR-1998; 98US-0083575.
XX
XX ('TAMO-') TAMOX INC.
XX
XX N1 B, Sun BNC, Sun CRV;
XX
XX WPI; 2000-052805/04.
XX
XX Treatment of neutropenia by stimulating proliferation of neutrophilic
XX cell lineage progenitors -
XX
XX Claim 13; Page 30; 64pp; English.
XX
XX The present sequence represents complementarity determining region
XX 2 (CDR2) of the light chain variable region of murine monoclonal
XX antibody Mab163-93. This antibody is an example of an agonist
XX molecule that specifically binds to or interacts with human
XX granulocyte colony stimulating factor (G-CSF) receptor to stimulate
XX cell proliferation and differentiation, especially by dimerising
XX the receptor or activating phosphorylation of kinases associated
XX with the receptor. Agonist antibodies can be used to stimulate
XX proliferation of G-CSF-dependent cells, e.g. to differentiate
XX leading to a repopulation of neutrophilic granulocyte lineage
XX cells, especially to treat neutropenia (claimed). They can also
XX be used to detect human G-CSF receptor immunologically.
XX
XX Sequence 7 AA:
XX
XX Query Match 100.0%; Score 38; DB 21; Length 7;
XX Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX QY 1 WASTRES 7
XX |||||
XX 1 WASTRES 7
XX
XX Db
XX
XX RESULT 9
XX AAE29275
XX ID AAE29275 standard; peptide; 7 AA.
XX
XX AC AAE29275;
XX
XX DT 27-JAN-2003 (first entry)
XX
XX DE Anti-ClfA monoclonal antibody variable light chain CDR2.
XX
XX KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
XX immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
XX septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy;
XX complementarity determining region; CDR.
XX
XX OS Unidentified.
XX
XX PN MO200272600-A2.
XX
XX PD 19-SEP-2002.
XX
XX PF 28-JAN-2002; 2002MO-US02296.
XX

```

XX	(INH-) INHIBITEX INC.
XX	
PI	Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
DR	WPI; 2002-759834/82.
XX	
PT	New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT	treating or preventing Staphylococcus aureus infection e.g. wound
PT	infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis
XX	in a human or animal
XX	
PS	Claim 30; Page 55; 80pp: English.
XX	
CC	The invention relates to monoclonal antibody which binds the clumping
CC	factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
CC	monoclonal antibody is useful for treating or preventing S. aureus
CC	infection in a human or animal, and for inhibiting the binding of
CC	staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC	S. aureus ClfA6 protein, S. aureus ClfB3 protein, or the S. aureus N3
CC	protein is useful for inducing an immunological response in a human or
CC	animal. These staphylococcal infections include wound infections, sepsis,
CC	impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC	present sequence is an anti-ClfA monoclonal antibody variable light chain
CC	complementarity determining region (CDR).
XX	
SO	Sequence 7 AA:
Query Match	100.0%; Score 38; DB 23; Length 7;
Best Local Similarity	100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 WASTRES 7 1 WASTRES 7
Db	
RESULT 10	
ID	ABP62361
XX	ABP62361 standard; Peptide; 7 AA.
AC	
XX	ABP62361;
DT	10-OCT-2002 (first entry)
DE	Human immunopeptide to HCV E2 glycoprotein light chain CDR #38.
XX	
KW	Viruslike: human: immunopolypeptide: immunopeptide: envelope glycoprotein;
KW	nonstructural protein; hepatitis C virus; HCV; E2 glycoprotein;
KW	N33 protein; viral infection.
OS	Homo sapiens.
XX	
PN	WO200259340-A1.
XX	
PD	01-AUG-2002.
XX	
PF	25-JAN-2002; 2002WO-US02303.
XX	
PR	26-JAN-2001; 2001US-264451P.
XX	
PA	(SCRI) SCRIPPS RES INST.
XX	
PI	Mariyama T, Jones IM, Burton DR, Fox RI;
XX	
DR	WPI; 2002-599801/64.
XX	
PT	New human immunopolypeptide with binding specificity for certain
PT	envelope glycoproteins and nonstructural proteins of hepatitis C virus

PT (HCV), for diagnosing or treating patients having or suspected of
 having HCV infection
 XX
 PS Claim 1; Fig 17; 308pp; English.
 CC The present invention relates to human immunopolypeptides, produced by a
 CC plaque transfected cell library. The present sequence is one such
 CC immunopolypeptide. The immunopolypeptides have binding specificity for
 CC envelope glycoprotein E2 and nonstructural protein NS3 of hepatitis C
 CC virus (HCV). E2 glycoprotein is believed to be responsible for target
 CC cell binding and contains neutralising epitopes, while NS3 is thought to
 CC be involved in the replication of HCV. The immunopolypeptides are useful
 CC for diagnosing and treating a patient having or suspected to be having
 CC HCV infection.
 CC
 SQ Sequence 7 AA;
 XX
 Query Match 100.0%; Score 38; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 WASTRES 7
 DB 1 WASTRES 7
 RESULT 11
 ID ABG68854 standard; Peptide: 7 AA.
 AC
 XX ABG68854;
 AC
 XX 07-OCT-2002 (first entry)
 DT
 XX CDR2 VL peptide.
 DE
 XX
 XX Cytostatic; virucide; hepatotropic; antinflammatory; neuroprotective;
 KM immunosuppressive; antiarthritic; cytokine receptor; interferon; IFN;
 KM cancer; haematological malignancy; viral infection; hepatitis; human;
 KM multiple sclerosis; autoimmune disease; arthritis.
 XX
 OS Synthetic.
 XX
 PN WO200244197-A2.
 PD
 XX 06-JUN-2002.
 PF
 XX 30-NOV-2001; 2001WO-CA01701.
 PR
 XX 01-DEC-2000; 2000US-0727388.
 PA (FISH/) FISH E. N.
 PI
 XX Fish EN;
 DR WPI; 2002-547689/58.
 XX N-PSDB; ABK37817.
 PT
 XX Cytokine receptor binding peptide construct, in particular interferon
 PT receptor binding peptide construct for use as an interferon mimetic,
 PT comprises a cytokine receptor binding domain incorporated in a
 PT molecular scaffold
 PS
 XX Example 8; Page 50; 105pp; English.
 CC This invention relates to a novel cytokine receptor binding peptide
 CC construct comprising a cytokine receptor binding domain incorporated in
 CC a suitable molecular scaffold so that the scaffold maintains the binding
 CC domain in a configuration suitable for binding to the cytokine receptor.
 CC The peptides of the invention may have cytostatic, virucide,
 CC hepatotropic, antinflammatory, neuroprotective, immunosuppressive and
 CC antiarthritic activities. A new interferon receptor binding peptide and
 CC construct is useful in the manufacture of a medicament as an interferon

CC (IFN) mimetic. A peptide that mimics the effect of IFN is useful in
 CC medical therapies for cancer, haematological malignancies, viral
 CC infections (hepatitis B or C), multiple sclerosis and autoimmune
 CC diseases such as arthritis, to detect modulators of IFN action, in
 CC screening assays to compare the activity and/or interaction with
 CC another molecule or potential IFN modulator and also in the diagnosis
 CC of IFN activity related disorders. A nucleic acid encoding the
 CC peptide of the invention or its useful for the treatment and therapy of
 CC the mentioned medical conditions. The peptide of the invention has less
 CC side effect than those of native cytokines. The present sequence
 CC represents an interferon receptor binding peptide of the invention.
 CC
 SQ Sequence 7 AA;
 XX
 Query Match 100.0%; Score 38; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 WASTRES 7
 DB 1 WASTRES 7
 RESULT 12
 ID AA081257 standard; Peptide: 7 AA.
 AC
 XX AA081257;
 AC
 XX 09-APR-2002 (first entry)
 DT
 XX Human trkC antibody light chain CDR2 of variable region #2.
 DE
 XX
 XX Human; mouse; anti-trkC agonist monoclonal antibody; trkC; antibody;
 KM trkA; trkB; cisplatin-induced neuropathy; pyridoxine-induced neuropathy;
 KM peripheral neuropathy; diabetic neuropathy; neurodegenerative disease;
 KM large-fibre sensory neuropathy; amyotrophic lateral sclerosis; tumour;
 KM nerve cell injury; blood cell disorder; leukopenia; eosinopenia; wound;
 KM basopenia; lymphopenia; monocytopenia; neutropenia; cancer; ulcer;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
 KM sickle cell disease; cardiac ischaemia; cerebrovascular disorder;
 KM cellular degeneration; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200198361-A2.
 PD
 XX 27-DEC-2001.
 PF
 XX 22-JUN-2001; 2001WO-US20153.
 PR
 XX 22-JUN-2000; 2000US-213141P.
 PR 05-OCT-2000; 2000US-238319P.
 PA (GENE) GENENTECH INC.
 PI
 XX Devaux B, Hongo JS, Presta LG, Shelton DL;
 DR WPI; 2002-130790/17.
 PT
 XX Novel anti-trkC agonist monoclonal antibody useful for treating
 PT neurodegenerative disease, shows no significant cross-reactivity with
 PT trkA/trkB, and recognizes epitope in domain 5 of trkC
 PS
 XX Claim 13; Fig 11; 121pp; English.
 CC The invention relates to an anti-trkC agonist monoclonal antibody which
 CC shows no significant cross-reactivity with trkA or trkB, and recognizes
 CC an epitope in domain 5 of trkC. The antibodies of the invention are
 CC effective in the treatment of cisplatin- or pyridoxine-induced
 CC neuropathy, peripheral neuropathy, diabetic neuropathy and large-fibre
 CC sensory neuropathy, neurodegenerative disease including amyotrophic
 CC lateral sclerosis, nerve cell injuries, disorders of insufficient blood

Query Match 100.0%; Score 38; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
 |||||
 Db 1 WASTRES 7

RESULT 15

ABP98681
 ID ABP98681 standard; peptide; 7 AA.

XX AC ABP98681;

XX DT 30-MAY-2003 (first entry)

XX DE Human GAH antibody light chain hypervariable region #2.

XX KW human; antibody; epitope; cytostatic; diagnosis; fluorescence wavelength;
 XX KW GAH antibody light chain; hypervariable region; cancer therapy.

XX OS Homo sapiens.

XX PN WO2003010542-A1.

XX PD 06-FEB-2003.

XX PF 25-JUL-2002; 2002WO-JP07547.

XX PR 25-JUL-2001; 2001JP-0224054.

XX PA (MITS-) MITSUBISHI PHARMA CORP.

XX PI Niki H, Tagawa T, Hosokawa S;

XX DR WPI; 2003-229648/22.

PT Cancer diagnostics based on peptide or protein recognizing part of
 PT tissue slice for staining and detection, also applicable in selecting
 PT remedies for cancer therapy and analysis of expression or behaviour of
 PT protein -

XX PS Claim 8; Page 46; 52pp; Japanese.

XX CC The invention relates to diagnostic agents comprising a peptide or a
 CC protein capable of recognizing at least a part of a tissue slice and
 CC a fluorescent substance characterized by: (a) having a fluorescence
 CC wavelength not close to the wavelength region of a non-specific
 CC autofluorescence of the tissue slice at a definite excitation wavelength;
 CC and (b) allowing simultaneous observation of image of the peptide or
 CC protein and that of the tissue slice. This peptide is an example of
 CC a peptide used in the method of the invention and represents the
 CC hypervariable region #2 from the human GAH antibody light chain.
 CC The diagnostic agents are for the detection of cancer, which are also
 CC applicable in selecting remedies for cancer therapy and analysis of
 CC expression and/or behaviour of proteins.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 24; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
 |||||
 Db 1 WASTRES 7

Search completed: October 7, 2003, 19:13:53
 Job time : 9.81034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 3.62069 Seconds
(without alignments)
305,878 Million cell updates/sec

Title: US-09-988-013a-2_COPY_56_62

Perfect score: 38

Sequence: 1 WASTRES 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	7	10 US-09-828-708-43	Sequence 43, App1
2	38	100.0	7	10 US-09-828-708-46	Sequence 46, App1
3	38	100.0	7	10 US-09-249-011A-18	Sequence 18, App1
4	38	100.0	7	11 US-09-563-222-15	Sequence 15, App1
5	38	100.0	7	11 US-09-563-222-27	Sequence 27, App1
6	38	100.0	7	12 US-10-071-962-19	Sequence 19, App1
7	38	100.0	7	14 US-10-146-305-10	Sequence 10, App1
8	38	100.0	82	15 US-10-078-958-15	Sequence 15, App1
9	38	100.0	94	9 US-09-905-243-60	Sequence 60, App1
10	38	100.0	101	14 US-10-025-687-25	Sequence 25, App1
11	38	100.0	101	15 US-10-194-975-89	Sequence 89, App1
12	38	100.0	101	15 US-10-125-687-25	Sequence 25, App1
13	38	100.0	107	12 US-10-325-694-142	Sequence 142, App1
14	38	100.0	107	12 US-10-325-694-148	Sequence 148, App1
15	38	100.0	107	12 US-10-325-694-149	Sequence 149, App1

16	38	100.0	107	12 US-10-325-694-152	Sequence 152, App1
17	38	100.0	109	10 US-09-828-708-1	Sequence 1, App1
18	38	100.0	109	10 US-09-828-708-4	Sequence 4, App1
19	38	100.0	112	12 US-10-229-335-4	Sequence 4, App1
20	38	100.0	112	12 US-10-229-335-28	Sequence 28, App1
21	38	100.0	112	15 US-10-056-052-6	Sequence 6, App1
22	38	100.0	112	15 US-10-056-052-10	Sequence 10, App1
23	38	100.0	112	15 US-10-056-052-14	Sequence 14, App1
24	38	100.0	112	15 US-10-056-052-18	Sequence 18, App1
25	38	100.0	113	10 US-09-741-843-2	Sequence 2, App1
26	38	100.0	113	10 US-09-741-843-6	Sequence 6, App1
27	38	100.0	113	10 US-09-274-1635-16	Sequence 16, App1
28	38	100.0	113	10 US-09-956-206A-80	Sequence 80, App1
29	38	100.0	113	11 US-09-894-839-2	Sequence 2, App1
30	38	100.0	113	11 US-09-894-839-6	Sequence 6, App1
31	38	100.0	113	11 US-09-894-839-20	Sequence 20, App1
32	38	100.0	113	11 US-09-895-529-6	Sequence 6, App1
33	38	100.0	113	11 US-09-988-013A-2	Sequence 2, App1
34	38	100.0	113	11 US-09-988-013A-6	Sequence 6, App1
35	38	100.0	113	12 US-10-159-006-2	Sequence 2, App1
36	38	100.0	113	12 US-10-159-006-4	Sequence 4, App1
37	38	100.0	113	12 US-10-159-006-6	Sequence 6, App1
38	38	100.0	113	12 US-10-159-006-32	Sequence 32, App1
39	38	100.0	113	12 US-10-159-006-33	Sequence 33, App1
40	38	100.0	113	12 US-10-159-006-34	Sequence 34, App1
41	38	100.0	113	12 US-10-330-613-22	Sequence 22, App1
42	38	100.0	113	12 US-10-330-530-22	Sequence 22, App1
43	38	100.0	113	12 US-10-160-506-71	Sequence 71, App1
44	38	100.0	113	12 US-10-160-506-91	Sequence 91, App1
45	38	100.0	113	12 US-09-215-163-42	Sequence 42, App1

ALIGNMENTS

RESULT 1
US-09-828-708-43
; Sequence 43, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their part
; FILE REFERENCE: 1361.005US1
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-708-43

Query Match 100.0%; Score 38; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
|||||
Db 1 WASTRES 7

RESULT 2
US-09-828-708-46
; Sequence 46, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.

```

; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their particl
; TITLE OF INVENTION: autoimmune disease
; FILE REFERENCE: 1361.005051
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-708-46

Query Match
Best Local Similarity 100.0%; Score 38; DB 10; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 1 WASTRES 7

RESULT 3
US-09-249-011A-18
; Sequence 18, Application US/09249011A
; Patent No. US20020176855A1
; GENERAL INFORMATION:
; APPLICANT: CO. MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; APPLICANT: CARRENO, BEATRIZ
; APPLICANT: CELINKER, ABBIE CHERYL
; APPLICANT: COLLINS, MARY
; APPLICANT: GOLDMAN, SAMUEL
; APPLICANT: GRAY, GARY S.
; APPLICANT: KNIGHT, ANDREA
; APPLICANT: O'HARA, DENISE
; APPLICANT: RUP, BONITA
; APPLICANT: VELDMAN, GERTRUUDA M.
; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
; TITLE OF INVENTION: OF TREATMENT THEREWITH
; FILE REFERENCE: 08702.0081-00000
; CURRENT APPLICATION NUMBER: US/09/249,011A
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CDR2 of humanized
US-09-249-011A-18

Query Match
Best Local Similarity 100.0%; Score 38; DB 10; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 1 WASTRES 7

RESULT 4
US-09-563-222-15
; Sequence 15, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hlatc, Andrew
; APPLICANT: Hehn, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; TITLE OF INVENTION: EUKARYOTIC CELLS
; FILE REFERENCE: 310098.406

```

```

; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-563-222-15

Query Match
Best Local Similarity 100.0%; Score 38; DB 11; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 1 WASTRES 7

RESULT 5
US-09-563-222-27
; Sequence 27, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hlatc, Andrew
; APPLICANT: Hehn, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; FILE REFERENCE: 310098.406
; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-563-222-27

Query Match
Best Local Similarity 100.0%; Score 38; DB 11; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 1 WASTRES 7

RESULT 6
US-10-071-962-19
; Sequence 19, Application US/10071962
; Publication No. US20030170237A1
; GENERAL INFORMATION:
; APPLICANT: Baufu N1
; APPLICANT: Bill N.C. Sun
; APPLICANT: Cedilly R.Y. Sun
; TITLE OF INVENTION: G-CSF Receptor Agonist Antibodies and
; FILE REFERENCE: 98-3
; CURRENT APPLICATION NUMBER: US/10/071,962
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US/09/303,155A
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 60/083,575
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 7
; TYPE: PRT
; ORGANISM: mouse
US-10-071-962-19

```

Query Match 100.0%; Score 38; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 1 WASTRES 7

RESULT 7
US-10-146-305-10
; Sequence 10, Application US/10146305
; Publication No. US20020173035A1
; GENERAL INFORMATION:
; APPLICANT: YUHAN CORPORATION
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV
; TITLE OF INVENTION: S-SURFACE ANTIGEN AND A GENE ENCODING THE SAME
; FILE REFERENCE: OY17440
; CURRENT APPLICATION NUMBER: US/10/146,305
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: KR 10-2001-26634
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Kopatentlin 1.71
; SEQ ID NO 10
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-146-305-10

Query Match 100.0%; Score 38; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 1 WASTRES 7

RESULT 8
US-10-078-958-15
; Sequence 15, Application US/10078958
; Publication No. US20030070185A1
; GENERAL INFORMATION:
; APPLICANT: JAKOBOVITS, AYA
; APPLICANT: KUCHERLAPATI, RAJU
; APPLICANT: KLAPHOLZ, SUSAN
; APPLICANT: MENDEZ, MICHAEL J.
; APPLICANT: GREEN, LARRY
; TITLE OF INVENTION: TRANSGENIC MAMMALS HAVING HUMAN Ig LOCI INCLUDING
; TITLE OF INVENTION: PLURAL VH AND VK REGIONS AND ANTIBODIES PRODUCED
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: CELL 4.18 CON
; CURRENT APPLICATION NUMBER: US/10/078,958
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 08/7759,620
; PRIOR FILING DATE: 1996-12-03
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-078-958-15

Query Match 100.0%; Score 38; DB 15; Length 82;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 37 WASTRES 43

RESULT 9
US-09-905-243-60
; Sequence 60, Application US/09905243
; Patent No. US20020062009A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Alexander H
; TITLE OF INVENTION: Monoclonal Antibodies with Reduced
; TITLE OF INVENTION: Immunogenicity
; FILE REFERENCE: P50770
; CURRENT APPLICATION NUMBER: US/09/905,243
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/300,970
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Macaca cynomolgus
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (24)...(40)
; OTHER INFORMATION: CDRI
; NAME/KEY: DOMAIN
; LOCATION: (56)...(62)
; OTHER INFORMATION: CDRII
US-09-905-243-60

Query Match 100.0%; Score 38; DB 9; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 56 WASTRES 62

RESULT 10
US-10-025-687-25
; Sequence 25, Application US/10025687
; Publication No. US20020142235A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Peter
; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
; FILE REFERENCE: 26050-705
; CURRENT APPLICATION NUMBER: US/10/025,687
; CURRENT FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentlin version 3.1
; SEQ ID NO 25
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-687-25

Query Match 100.0%; Score 38; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 56 WASTRES 62

RESULT 11
US-10-194-975-89
; Sequence 89, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01

```

; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-89

Query Match
Best Local Similarity 100.0%; Score 38; DB 15; Length 101;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
   |||||
Db 56 WASTRES 62

RESULT 12
US-10-125-687-25
; Sequence 25, Application US/10125687
; Publication No. US20030054407A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Peter
; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
; FILE REFERENCE: 26050-705
; CURRENT APPLICATION NUMBER: US/10/125,687
; CURRENT FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-687-25

Query Match
Best Local Similarity 100.0%; Score 38; DB 15; Length 101;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
   |||||
Db 56 WASTRES 62

RESULT 13
US-10-325-694-142
; Sequence 142, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 142
; LENGTH: 107
; TYPE: PRT
; ORGANISM: HUMAN
US-10-325-694-142

Query Match
Best Local Similarity 100.0%; Score 38; DB 12; Length 107;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WASTRES 7
   |||||
Db 50 WASTRES 56

RESULT 14
US-10-325-694-148
; Sequence 148, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 148
; LENGTH: 107
; TYPE: PRT
; ORGANISM: HUMAN
US-10-325-694-148

Query Match
Best Local Similarity 100.0%; Score 38; DB 12; Length 107;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
   |||||
Db 50 WASTRES 56

RESULT 15
US-10-325-694-149
; Sequence 149, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 107
; TYPE: PRT
; ORGANISM: HUMAN
US-10-325-694-149

Query Match
Best Local Similarity 100.0%; Score 38; DB 12; Length 107;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
   |||||
Db 50 WASTRES 56

Search completed: October 7, 2003, 19:24:24
Job time : 3.62069 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 2.6069 Seconds

(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013a-2_COPY_56_62

Perfect score: 38

Sequence: 1 WASTRES 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:***
2: PIR1:***
3: PIR2:***
4: PIR3:***
5: PIR4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	81	2 PH1048	Ig light chain V r
2	38	100.0	92	2 S37533	Ig kappa chain V r
3	38	100.0	92	2 S37530	Ig kappa chain V r
4	38	100.0	92	2 S37529	Ig kappa chain V r
5	38	100.0	92	2 S37535	Ig kappa chain V r
6	38	100.0	92	2 S37531	Ig kappa chain V r
7	38	100.0	92	2 S37534	Ig kappa chain V r
8	38	100.0	92	2 S37532	Ig kappa chain V r
9	38	100.0	98	2 S19974	Ig kappa chain V r
10	38	100.0	101	2 S26337	Ig kappa chain V r
11	38	100.0	101	2 PH1046	Ig light chain V r
12	38	100.0	101	2 PH0869	Ig kappa chain V r
13	38	100.0	102	2 A34153	Ig kappa chain V-I
14	38	100.0	103	2 PH1050	Ig light chain V r
15	38	100.0	103	2 PH1051	Ig light chain V r
16	38	100.0	103	2 PH1052	Ig light chain V r
17	38	100.0	103	2 PH1054	Ig light chain V r
18	38	100.0	103	2 PH1047	Ig light chain V r
19	38	100.0	104	2 PH1101	Ig light chain V r
20	38	100.0	104	2 PH1102	Ig light chain V r
21	38	100.0	104	2 PH1103	Ig light chain V r
22	38	100.0	104	2 PH1104	Ig light chain V r
23	38	100.0	104	2 F38601	Ig kappa chain V r
24	38	100.0	106	2 A49138	IgA kappa rheumato
25	38	100.0	111	2 S03304	Ig kappa chain V r
26	38	100.0	111	2 G30502	Ig kappa chain V r
27	38	100.0	112	2 P10265	Ig kappa chain V r
28	38	100.0	112	2 S09970	Ig kappa chain V-J
29	38	100.0	112	2 S43103	Ig kappa chain V-J

30	38	100.0	113	2 S30523	Ig kappa chain V r
31	38	100.0	113	2 P10263	Ig kappa chain V r
32	38	100.0	113	2 S34003	Ig kappa chain V r
33	38	100.0	113	2 S34002	Ig kappa chain V r
34	38	100.0	114	1 K4HUN	Ig kappa chain V-I
35	38	100.0	114	2 S44116	Ig kappa chain V-J
36	38	100.0	114	2 S44119	Ig kappa chain V-J
37	38	100.0	118	2 PT0356	Ig kappa chain V r
38	38	100.0	120	2 G33932	Ig kappa chain pre
39	38	100.0	120	2 S51147	Ig kappa chain pre
40	38	100.0	121	1 K4HU	antibody light cha
41	38	100.0	129	2 S40329	Ig kappa chain pre
42	38	100.0	129	2 S40347	Ig kappa chain V-J
43	38	100.0	133	1 K4HUI	Ig kappa chain pre
44	38	100.0	133	2 PS0023	Ig kappa chain pre
45	38	100.0	134	2 PC1214	Ig kappa chain pre

ALIGNMENTS

RESULT 1

PH1048
Ig light chain V region (clone 165.49) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Jun-1996

C:Accession: PH1048

R:Tillman, D.M.; Jou, N.T.; Hall, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH1048

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-81 <Till>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: Immunoglobulin

Query Match Similarity 100.0%; Score 38; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WASTRES 7

Db 35 WASTRES 41

|||||

RESULT 2

S37533
Ig kappa chain V region (V-Kappa 4) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999

C:Accession: S37533

R:Klein, U.; Kuipers, R.; Rajewsky, K.

submitted to the EMBL Data Library, September 1993

A:Description: Human IgM(+)IgD(+) cells, the major B cell subset in the peripheral bl

A:Reference number: S37501

A:Accession: S37533

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-92 <Klein>

A:Cross-references: EMBL:Z26627; NID:9405708; PTDN:CAAB1380.1; PTD:9405709

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

Query Match Similarity 100.0%; Score 38; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 0.34; 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WASTRES 7

Db 34 WASTRES 40

|||||

```
RESULT 3
S37530
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37530
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37530
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:226630; NID:g405702; PIDN:CAA81383.1; PID:g405703
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 38; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 34 WASTRES 40

RESULT 4
S37529
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37529
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:226631; NID:g405700; PIDN:CAA81384.1; PID:g405701
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 38; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 34 WASTRES 40

RESULT 5
S37535
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37535
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37535
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:226625; NID:g405712; PIDN:CAA81378.1; PID:g405713
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 34 WASTRES 40

RESULT 6
S37531
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37531
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37531
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:226629; NID:g405704; PIDN:CAA81382.1; PID:g405705
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 38; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 34 WASTRES 40

RESULT 7
S37534
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37534
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37534
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:226626; NID:g405710; PIDN:CAA81379.1; PID:g405711
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 38; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 34 WASTRES 40

RESULT 8
S37532
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37532
R:Klein, U.; Kueppers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IDP(+) cells, the major B cell subset in the peripheral blood
```


A:Reference number: S37501
A:Accession: S37532
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <RLE>
A:Cross-references: EMBL:Z26628; NID:9405706; PIDN:CAA81381.1; PID:9405707
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 38; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
|||||||
DB 34 WASTRES 40

RESULT 9

S19974
Ig kappa chain V region (M-T406) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jan-2000
C:Accession: S19974
R:Weissenborn, W.; Riethmuelier, G.; Weiss, E.M.; Rieber, E.P.
A:Description: Structural characterization of CD4 mab.
A:Reference number: S19963
A:Accession: S19974
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-98 <WEI>
A:Cross-references: EMBL:X65096; NID:952294; PIDN:CAA46224.1; PID:952295
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-80/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
|||||||
DB 40 WASTRES 46

RESULT 10

S26337
Ig light chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
C:Accession: S26337; S78449
R:Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A:Title: Antibodies that are specific for a single amino acid interchange in a protein
A:Reference number: S26309; MUID:91341421; PMID:1908510
A:Accession: S26337
A:Molecule type: mRNA
A:Residues: 1-101 <SFA>
A:Cross-references: EMBL:X59193
R:Caton, A.J.
submitted to the EMBL Data Library, April 1991
A:Reference number: S78447
A:Accession: S78449
A:Molecule type: mRNA
A:Residues: 1-60, 'T', 62-91, 'S', 93-101 <CAT>
A:Cross-references: EMBL:X59193; NID:952323; PIDN:CAA1903.1; PID:91334067
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:8-88/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.37;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WASTRES 7
|||||||
DB 48 WASTRES 54

RESULT 11

PH1046
Ig light chain V region (clone 202.9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PH1046
R:Trillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0871; MUID:92381444; PMID:1512540
A:Accession: PH1046
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-101 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]P1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin
F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
|||||||
DB 56 WASTRES 62

RESULT 12

PH0869
Ig kappa chain V region (anti-DNA, H2F) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 21-Jan-2000
C:Accession: PH0869
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghosssein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
A:Reference number: PH0862; MUID:92078875; PMID:1660528
A:Accession: PH0869
A:Molecule type: DNA
A:Residues: 1-101 <MAN>
A:Note: residues 28-33 were obtained from Figure 4
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that be
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-23/Region: framework 1
F:16-96/Domain: Immunoglobulin homology <IMM>
F:24-40/Region: complementarity-determining 1
F:41-55/Region: framework 2
F:56-62/Region: complementarity-determining 2
F:63-94/Region: framework 3
F:95-101/Region: complementarity-determining 3

Query Match 100.0%; Score 38; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
|||||||
DB 56 WASTRES 62

RESULT 13

A34153
Ig kappa chain V-IV region (Dep) - human
C:Species: Homo sapiens (man)

C>Date: 11-Apr-1990 #sequence_revision 11-Apr-1990 #text_change 21-Jan-2000
C:Accession: A34153
R:Milnesco, E.; Ayadi, H.; Congy, N.; Gendton, M.C.; Roy, J.P.; Heyermann, H.; Frangione
J. Biol. Chem. 264, 21481-21485, 1989
A:Title: Multiple mutations in the variable region of the kappa light chains of three mo
A:Reference number: A34153; MUID:90094313; PMID:2480953
A:Accession: A34153
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-102 <MIH>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
|||||
DB 56 WASTRES 62

RESULT 14

PH1050
Ig light chain V region (clone 111-c1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PH1050
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c
A:Reference number: PH0971; MUID:92381444; PMID:1512540
A:Accession: PH1050
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-103 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
|||||
DB 56 WASTRES 62

RESULT 15

PH1051
Ig light chain V region (clone 165.3) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PH1051
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c
A:Reference number: PH0971; MUID:92381444; PMID:1512540
A:Accession: PH1051
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-103 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
|||||
DB 56 WASTRES 62

Search completed: October 7, 2003, 19:21:47
Job time : 2.6069 secs

Gencore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 1.42414 Seconds
(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-2_COPY_56_62
Perfect score: 38
Sequence: 1 WASTRES 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	109	RV4D_HUMAN	P83593 homo sapien
2	38	100.0	114	RV4A_HUMAN	P01625 homo sapien
3	38	100.0	121	RV4B_HUMAN	P06312 homo sapien
4	38	100.0	133	RV4C_HUMAN	P06313 homo sapien
5	38	100.0	134	RV4C_HUMAN	P06314 homo sapien
6	32	84.2	685	SYM_NEIM	O91790 neisseria m
7	31	81.6	382	THRC_SYNY3	P74193 synecocyst
8	30	78.9	497	GLPK_PYRKO	O93623 pyrococcus
9	30	78.9	668	COAT_FCVF4	P27405 feline cali
10	30	78.9	677	SG1_RAT	O35314 rattus norv
11	30	78.9	677	SG1_MOUSE	O35314 rattus norv
12	30	78.9	1139	HMW1_MOUSE	O49413 mycoplasma
13	30	78.9	2365	CCAH_MOUSE	O49413 mycoplasma
14	30	78.9	3414	POLG_TBRYH	P01299 t genome po
15	30	78.9	3414	POLG_TBRYH	P01299 t genome po
16	29	76.3	227	ATP6_BRAFL	O47426 branchiosto
17	29	76.3	227	ATP6_BRAFL	O47426 branchiosto
18	29	76.3	413	RCA_CUCSA	O01587 cucumis sat
19	29	76.3	420	O63A_DROME	O92528 drosophila
20	29	76.3	425	RCA_B_HORYU	O42450 hordeum vul
21	29	76.3	430	HEMA_CVMS	P31614 murine coro
22	29	76.3	433	RCA_MALIZ	O92100 zea mays (m
23	29	76.3	437	RCA_MALIZ	O92100 zea mays (m
24	29	76.3	439	HEMA_CVMS	O83356 murine domes
25	29	76.3	439	HEMA_CVMS	O83356 murine domes
26	29	76.3	441	RCA_TOBAC	O40565 nicotiana t
27	29	76.3	442	RCA_TOBAC	O40565 nicotiana t
28	29	76.3	451	RCA_TOBAC	O40460 phaseolus v
29	29	76.3	459	RCA_LYCPN	O93514 homo sapien
30	29	76.3	472	RCA_SPIOL	O49074 lycopersico
31	29	76.3	474	RCA_ABYTH	P10871 spinaclia ol
32	29	76.3	574	MP12_RAT	P10896 arabisopsis
33	29	76.3	580	MP12_HUMAN	P84966 rattus norv
					P30305 homo sapien

34	29	76.3	685	1	SYM_NEIM	O91790 neisseria m
35	29	76.3	717	1	MCCA_MOUSE	O99m8 mus musculu
36	29	76.3	1653	1	YFHM_ECOLI	P76578 escherichia
37	28	73.7	79	1	Y524_MYCLE	O05755 mycobacteri
38	28	73.7	179	1	KPFA_FUSNN	O8557 fusobacteri
39	28	73.7	203	1	SP24_MOUSE	O8113 mus musculu
40	28	73.7	204	1	VAS_VESFL	P35783 vesipula fla
41	28	73.7	204	1	VAS_VESGE	P35784 vesipula ger
42	28	73.7	204	1	VAS_VESMC	P35760 vesipula mac
43	28	73.7	205	1	VAS_VESPE	P35785 vesipula pen
44	28	73.7	205	1	RT26_HUMAN	O92528 homo sapien
45	28	73.7	217	1	TAL_CAUCR	O92521 caulobacter

ALIGNMENTS

```

RESULT 1
RV4D_HUMAN          STANDARD:      PRT:      109 AA.
ID RV4D_HUMAN
AC P83593;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 19 kappa chain V-IV region STR (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Abdominal adipose tissue;
RA PubMed-9588180;
RT Olsen K.E., Sletten K., Westermarck P.;
RT "Extended analysis of Al-amyloid protein from abdominal wall
RT subcutaneous fat biopsy: kappa IV immunoglobulin light chain.";
RT Blochem. Biophys. Res. Commun. 245:713-716(1998).
CC -1 - FUNCTION: May play an important role in fibrillogenesis.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003599; IG.
DR InterPro: IPR003596; IG-V.
DR Pfam: PF00047; 19; 1.
DR PROSITE: PS50835; IG-LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 40 FRAMEWORK-1.
FT DOMAIN 41 55 FRAMEWORK-2.
FT DOMAIN 56 62 FRAMEWORK-2.
FT DOMAIN 63 94 FRAMEWORK-3.
FT DOMAIN 95 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 109 COMPLEMENTARITY-DETERMINING-4.
FT DISULFD 23 94 BY SIMILARITY.
FT UNSURE 23 23
FT UNSURE 94 94
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12060 MW; 0C4F31EA11E12A0B CRC64;

Query Match 100.0%; Score 38; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 56 WASTRES 62

RESULT 2
RV4A_HUMAN          STANDARD:      PRT:      114 AA.
ID RV4A_HUMAN
AC P01625;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

```

DE Ig kappa chain V-IV region Len.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=76004342; PubMed=50995;
RA Schneider M., Hilschmann N.;
RT "The primary structure of a monoclonic immunoglobulin-L-chain of
RT subgroup IV of the kappa type (Bence-Jones protein Len).";
RL Hoppe-Sejler's Z. Physiol. Chem. 356:507-557(1975).
RN [2]
RP REVISION TO 9.
RA Salomon A.;
RC Submitted (AUG-1996) to the SWISS-PROT data bank.
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
DR PDB: 1LEO; 01-FEB-01.
DR PDB: 1LEU; 03-FEB-01.
DR PDB: 1ERQ; 09-FEB-01.
DR PDB: 1EK3; 06-MAR-01.
DR PDB: 1LVE; 21-JAN-98.
DR PDB: 1LVE; 18-MAY-99.
DR PDB: 1LVE; 28-MAR-01.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR pfam: PF00047; Ig_V.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Bence-Jones protein; 3D-structure.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 40 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 41 55 FRAMEWORK-2.
FT DOMAIN 56 62 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 63 94 FRAMEWORK-3.
FT DOMAIN 95 101 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 102 113 FRAMEWORK-4.
FT NON_TER 114 114 BY SIMILARITY.
SQ SEQUENCE 114 AA; 12640 MW; 0647FD17F236485 CRC64;

Query Match 100.0%; Score 38; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 56 WASTRES 62

RESULT 3
KV40_HUMAN STANDARD; PRT; 121 AA.
AC P06312;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig kappa chain V-IV region precursor (Fragment).
GN IGKV4-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86041853; PubMed=2997712;
RA Klobeck H.G., Bornkamm G.W., Combrato G., Mocklat R., Pohlenz H.D.,
RA Zachau H.G.;

```

```

RT Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene."
RL Nucleic Acids Res. 13:6515-6529(1985).
CC -1- MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 200023; CAA77318.1; -
DR PIR; A01902; K4HU.
DR HSSP; P80362; 1WTL.
DR GeneW; HGNC:5834; IGKV4-1.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR pfam: PF00047; IGV; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT DOMAIN 21 >121 IG KAPPA CHAIN V-IV REGION.
FT DOMAIN 44 60 FRAMEWORK-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 83 114 COMPLEMENTARITY-DETERMINING-2.
FT DISULFID 115 121 FRAMEWORK-3.
FT NON_TER 121 121 BY SIMILARITY.
SQ SEQUENCE 121 AA; 13380 MW; 9586AD4188D33974 CRC64;

Query Match 100.0%; Score 38; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 76 WASTRES 82

RESULT 4
KV4B_HUMAN STANDARD; PRT; 133 AA.
AC P06313;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region J1 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86041853; PubMed=2997712;
RA Klobeck H.G., Bornkamm G.W., Combrato G., Mocklat R., Pohlenz H.D.,
RA Zachau H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene."
RL Nucleic Acids Res. 13:6515-6529(1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: Z00022; CAA77317.1; -

DR PIR: A01904; K4HUIJ.

DR HSSP: P80362; 1MTL.

DR GO: GO:0005576; C:extracellular; NAS.

DR GO: GO:0003823; F:antigen binding activity; NAS.

DR GO: GO:0006955; P:immune response; NAS.

DR InterPro: IPR007110; Ig-like.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig; 1.

DR SMART: SM00406; IGV; 1.

DR PROSITE: PS50835; IG-LIKE; 1.

DR Immunoglobulin V region; Signal.

FT SIGNAL 1 20

FT CHAIN 1 133 IG KAPPA CHAIN V-IV REGION JT.

FT DOMAIN 21 43 FRAMEWORK-1.

FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 61 75 FRAMEWORK-2.

FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 83 114 FRAMEWORK-3.

FT DOMAIN 115 122 COMPLEMENTARITY-DETERMINING-3.

FT DISULFID 123 132 FRAMEWORK-4.

FT DISULFID 133 134 BY SIMILARITY.

FT NON_TER 133 133

SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;

Query Match 100.0%; Score 38; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 76 WASTRES 82

RESULT 5
KVAC_HUMAN STANDARD; PRT; 134 AA.

ID P06314;
AC 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 08, Last annotation update)
DE Ig kappa chain V-IV region B17 precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86041854; PubMed-2997713;
RA Marsh P., Mills F., Gould H.;
RT "Detection of a unique human V kappa IV germline gene by a cloned cDNA probe."
RL Nucleic Acids Res. 13:6531-6544(1985).
RN [2]
RP REVISION TO 76.
RA Marsh P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X02990; CAA26733.1; -
DR HSSP: P80362; 1MTL.

DR GO: GO:0005576; C:extracellular; NAS.

DR GO: GO:0003823; F:antigen binding activity; NAS.

DR GO: GO:0006955; P:immune response; NAS.

DR InterPro: IPR007110; Ig-like.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig; 1.

DR SMART: SM00406; IGV; 1.

DR PROSITE: PS50835; IG-LIKE; 1.

DR Immunoglobulin V region; Signal.

FT SIGNAL 1 20

FT CHAIN 1 134 IG KAPPA CHAIN V-IV REGION B17.

FT DOMAIN 21 43 FRAMEWORK-1.

FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 61 75 FRAMEWORK-2.

FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 83 114 FRAMEWORK-3.

FT DOMAIN 115 121 FRAMEWORK-4.

FT DOMAIN 122 133 FRAMEWORK-4.

FT DISULFID 134 134 BY SIMILARITY.

SQ SEQUENCE 134 AA; 14966 MW; 6413A22ED0738832 CRC64;

Query Match 100.0%; Score 38; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRES 7
Db 76 WASTRES 82

RESULT 6
SYM_NEIMA STANDARD; PRT; 685 AA.

ID Q0JWP0;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (methionine--tRNA ligase) (metr).
GN METG OR NMA0275.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID:65699;
RN [1]
RP SEQUENCE FROM N.A.
RC SYRAIN-22491 / Serogroup A / Serotype 4A;
RX MEDLINE-20222556; PubMed-10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S., Jagers K., Leather S., Moulle S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491."
RL Nature 404:502-506(2000).
CC -1- FUNCTION: is required not only for elongation of protein synthesis but also for the initiation of all mRNA translation through initiator tRNA(fMet) aminoacylation (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) -> AMP + diphosphate + L-methionyl-tRNA(Met).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family. Metg subfamily 1.
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AL162752; CAB83583.1; -
CC PIR: G82022; G82022.
CC DR HSP: P00959; IMEX.
CC DR HAMAP; MF_00098; fused; 1.
CC DR InterPro; IPR004495; MetC_Cterm.
CC DR InterPro; IPR002300; trna-synt_1a.
CC DR InterPro; IPR001412; trna-synt_1.
CC DR InterPro; IPR002304; trna-synt_met.
CC DR Pfam; PF00133; trna-synt_1; 1.
CC DR Pfam; PF01588; trna_bind; 1.
CC DR PIRSF; PIRSF001528; MetRS_dimerising; 1.
CC DR PRINTS; PR01041; TRNASYNTHMET.
CC DR TIGRfam; TIGR00398; metC; 1.
CC DR TIGRfam; TIGR00399; metC_Cterm; 1.
CC DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
CC DR PROSITE; PS50886; TRBD; 1.
CC DR Aminoacyl-tRNA synthetase; Protein biosynthesis; ligase; ATP-binding;
CC RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
CC FT SITE 12 22 "HIGH" REGION.
CC FT SITE 339 343 "MASKS" REGION.
CC FT DOMAIN 582 685 TRNA-BINDING.
CC FT METAL 143 143 ZINC (BY SIMILARITY).
CC FT METAL 146 146 ZINC (BY SIMILARITY).
CC FT METAL 156 156 ZINC (BY SIMILARITY).
CC FT METAL 159 159 ZINC (BY SIMILARITY).
CC FT BINDING 342 342 ATP (BY SIMILARITY).
CC SQ SEQUENCE 685 AA; 76984 MW; 07FD45915ED3BEF3 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 685;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 520 WANTRET 526
11:1111

RESULT 7
THRC_SYNY3 STANDARD; PRT; 382 AA.
AC P74193;
DT 15-DEC-1998 (rel. 37; Created)
DT 15-DEC-1998 (rel. 37; Last sequence update)
DT 28-FEB-2003 (rel. 41; Last annotation update)
DE Threonine synthase (EC 4.2.3.1).
THRC OR SLI172.
OS *Synechocystis* sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905221;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugliura M., Sasanuma S., Kimura T.,
RA Hosoguchi T., Matsumoto A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT *Sequence analysis of the genome of the unicellular cyanobacterium
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Rep. 3:109-116(1996).
CC -1- CATALYTIC ACTIVITY: O-phospho-L-homoserine + H(2)O = L-threonine +
CC phosphate.
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Threonine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE SERINE/THREONINE DEHYDRATASE FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, D90913; BAA18282.1; -
CC PIR: S75823; S75823.
CC DR HSP: Q95785; 1ESX.
CC DR InterPro; IPR001926; B6_enzyme_beta.
CC DR InterPro; IPR000634; S/T_denhydrtse.
CC DR Pfam; PF00291; PALP; 1.
CC DR TIGRfam; TIGR00260; thrc; 1.
CC DR PROSITE; PS00165; DEHYDRATASE_SER_THR_1.
CC KW Threonine biosynthesis; Lyase; Pyridoxal phosphate; Complete proteome.
CC FT BINDING 93 93 PYRIDOXAL PHOSPHATE (PROBABLE).
CC SQ SEQUENCE 382 AA; 40414 MW; FBC9AEEC2B7F6F35 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 382;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WASTRES 7
DB 285 WASTRES 291
11:1111

RESULT 8
GLPK_PYRKO STANDARD; PRT; 497 AA.
ID GLPK_PYRKO
AC O93623;
DT 30-MAY-2000 (rel. 39; Created)
DT 30-MAY-2000 (rel. 39; Last sequence update)
DT 28-FEB-2003 (rel. 41; Last annotation update)
DE Glycerol kinase (EC 2.7.1.30) (ATP:glycerol 3-phosphotransferase)
DE (Glycerokinase) (GK).
DE GLPK.
OS *Pyrococcus kodakarensis*.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=99127861; PubMed=9930671;
RA Koga Y., Morikawa M., Haruki M., Nakamura H., Imanaka T., Kanaya S.;
RT "Thermotable glycerol kinase from a hyperthermophilic archaeon: gene
RT cloning and characterization of the recombinant enzyme.";
RL Protein Eng. 11:1219-1227(1998).
CC -1- FUNCTION: Key enzyme in the regulation of glycerol uptake and
CC metabolism. Can utilize other nucleoside triphosphates than ATP as
CC a phosphoryl donor.
CC -1- CATALYTIC ACTIVITY: ATP + glycerol = ADP + glycerol 3-phosphate.
CC -1- PATHWAY: Glycerol utilization; rate-limiting step.
CC -1- SUBUNIT: Homodimer.
CC -1- MISCELLANEOUS: The optimum pH and temperature are 8.0 and 80
CC degrees Celsius respectively.
CC -1- SIMILARITY: BELONGS TO THE GLUCOKINASE / GLUCONOKINASE /
CC GLYCEROKINASE / XYLOKINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AB012099; BAA34909.1; -

DR PIR: T43888; T43888.
 DR HAMAP: P08859; 1GLC.
 DR HAMAP: MF-00186; -; 1.
 DR InterPro: IPR005777; FGCV_kin.
 DR InterPro: IPR005999; Glycerol_kin.
 DR Pfam: PF02782; FGCV_1.
 DR TIGRfam: TIGR01311; glycerol_kin; 1.
 DR PROSITE: PS00445; FGCV_KINASES_2; 1.
 DR PROSITE: PS00933; FGCV_KINASES_1; 1.
 DR Glycerol metabolism; Transferase; Kinase; ATP-binding.
 FT NE_BIND 151 163 ATP (PROBABLY).
 SO SEQUENCE 497 AA; 55904 MW; 91A0A16D9735B56 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 497;
 Best Local Similarity 83.3%; Pred. No. 56;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WASTRE 6
 |||||
 DB 445 WADTRE 450

RESULT 9
 COAT_FCVF4 STANDARD; PRT; 668 AA.
 ID COAT_FCVF4
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 NC NCB1_Taxid=11980;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91306470; PubMed=1853578;
 RA Tshya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";
 RL Virology 183:810-814(1991).
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D90357; BAAL4371.1; -.
 DR PIR: B40481; VCMWRC.
 DR InterPro: IPR004005; Calic1_coat.
 DR Pfam: PF00915; Calic1_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 668 AA; 73589 MW; 85BDCB85804E503 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 668;
 Best Local Similarity 71.4%; Pred. No. 76;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WASTRES 7
 |||||
 DB 72 WASTHEA 78

RESULT 10
 SGL_RAT STANDARD; PRT; 675 AA.
 ID SGL_RAT
 AC 035314; 090YGB; 090VH1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Secretogranin I precursor (Sg1) (Chromogranin B) (Cgb) (Glucagonoma
 DE peptide) [Contains: GAWK peptide; CCB peptide].
 GN CGB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_Taxid=101116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-Sprague-Dawley;
 RX MEDLINE=90351888; PubMed=2641278;
 RA Forss-Petter S., Danielson P., Battenberg E., Bloom F.,
 RA Sutcliffe J.G.;
 RT "Nucleotide sequence and cellular distribution of rat chromogranin B
 RT (secretogranin I) mRNA in the neuroendocrine system.";
 RL J. Mol. Neurosci. 1:63-75(1989).

[2]
 PARTIAL SEQUENCE.
 RC TISSUE=Glucagonoma;
 RX MEDLINE=92063871; PubMed=1954895;
 RA Nielsen E., Welinder B.S., Madsen O.D.;
 RT "Chromogranin-B, a putative precursor of eight novel rat glucagonoma
 RT peptides through processing at mono-, di-, or tribasic residues.";
 RL Endocrinology 129:3147-3156(1991).
 CC -1- FUNCTION: NEUROENDOCRINE SECRETORY GRANULE PROTEIN, WHICH MAY BE
 CC THE PRECURSOR FOR OTHER BIOLOGICALLY ACTIVE PEPTIDES.
 CC -1- SUBCELLULAR LOCATION: NEUROENDOCRINE AND ENDOCRINE SECRETORY
 CC GRANULES (BY SIMILARITY).
 CC -1- PTM: EXTENSIVELY PROCESSED IN GLUCAGONOMA TISSUE BY LIMITED
 CC PROTEOLYSIS AT CONSERVED BASIC RESIDUES. ALTERNATIVE PROCESSING
 CC ARE SEEN IN DIFFERENT TISSUES. THE PROGLUCAGON-CONVERTING ENZYMES
 CC PRESENT IN TRANSFORMED ALPHA-CELLS ARE LIKELY CANDIDATES TO BE
 CC INVOLVED IN TISSUE-SPECIFIC PROCESSING.
 CC -1- SIMILARITY: BELONGS TO THE CHROMOGRANIN / SECRETOGRANIN PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF019974; AAB72089.1; -.
 DR InterPro: IPR001819; Chromogranin_AB.
 DR InterPro: IPR001900; Granin.
 DR Pfam: PF01271; Granin; 1.
 DR PRINTS: PR00659; CHROMOGRANIN.
 DR PROSITE: PS00422; GRANINS_1; 1.
 DR PROSITE: PS00423; GRANINS_2; 1.
 KW Sulfation; Cleavage on pair of basic residues; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 675
 FT PEPTIDE 435 508 SECRETOGRANIN I.
 FT PEPTIDE 615 674 GAWK PEPTIDE (BY SIMILARITY).
 FT DISULFID 36 57 CCB PEPTIDE (BY SIMILARITY).
 FT MOD_RES 171 171 BY SIMILARITY.
 FT MOD_RES 339 339 SULEFATION (POTENTIAL).
 SO SEQUENCE 675 AA; 77392 MW; 2D3D71D84C0ABE48 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 675;
 Best Local Similarity 83.3%; Pred. No. 77;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6
DB 97 WASSRE 102

RESULT 11

SG1_MOUSE STANDARD: PRT: 677 AA.

AC P16014;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Secretogranin I precursor (Sgi) (Chromogranin B) (CGB).
GN CHGB OR SGI1 OR SGI-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=pituitary;
RX MEDLINE=90206804; PubMed=2320426;
RA Linard C.G., Mikay M., Seldah N.G., Chretien M.;
RT "Primary structure of mouse chromogranin B deduced from cDNA
sequence.";
RL Nucleic Acids Res. 18:1298-1298(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90242932; PubMed=2335203;
RA Pohl T.M., Phillips E., Song K., Gerdes H.H., Huttner W.B.,
RA Ruehner U.;
RT "The organization of the mouse chromogranin B (secretogranin I)
gene.";
RL FEBS Lett. 262:219-224(1990).
CC -1- FUNCTION: SECRETOGRANIN I IS A NEUROENDOCRINE SECRETORY GRANULE
PROTEIN, WHICH MAY BE THE PRECURSOR FOR OTHER BIOLOGICALLY ACTIVE
PEPTIDES.
CC -1- SUBCELLULAR LOCATION: Neuroendocrine and endocrine secretory
granules.
CC -1- SIMILARITY: BELONGS TO THE CHROMOGRANIN / SECRETOGRANIN PROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X33028; CAA37199.1; -;
DR EMBL: X51429; CAA35792.1; -;
DR PIR: S09078; S09078.
DR MGI: 88395; Chgb.
DR InterPro: IPR001819; Chromogranin_AB.
DR InterPro: IPR001990; Granin.
DR Pfam: PFO1271; Granin_1.
DR PRINTS: PRO0659; CHROMOGRANIN.
DR PROSITE: PS00422; GRANIN_1; 1.
DR PROSITE: PS00423; GRANIN_2; 1.
KW Sulfation; Cleavage on pair of basic residues; Signal.
FT SIGNAL 1 20
FT CHAIN 21 677 SECRETOGRANIN I.
FT DISULFID 36 57 BY SIMILARITY.
FT MOD_RES 348 348 SULFATION (BY SIMILARITY).
FT CONFICT 422 424 REP -> IGA (IN REF. 1).
SO SEQUENCE 677 AA; 77969 MW; C7391E781E4F9B9 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 677;
Best Local Similarity 83.3%; Pred. No. 78;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6

DB 97 WASSRE 102

RESULT 12

HMW1_MYCGE STANDARD: PRT: 1139 AA.

AC Q49413; Q49365;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome high molecular weight protein 1 (cytochrome accessory
protein 1).
GN HMW1 OR MG312.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Frichman J.L., Weidman J.F., Small R.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort R.F., Hu P.-C., Luetter T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 721-847 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bort R.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHESIN PROTEINS
CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT
MEMBRANE (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U39712; AAC71534.1; -;
DR EMBL: U02261; AAD12527.1; -;
DR PIR: E64234; E64234.
DR TIGR: MG312; -;
KW Cytochrome; Structural protein; Complete proteome.
SO SEQUENCE 1139 AA; 130531 MW; 0011D3286C3DD856 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 1139;
Best Local Similarity 71.4%; Pred. No. 14+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 123 WISTRES 129

RESULT 13

CCAH_MOUSE STANDARD: PRT: 2365 AA.

AC 088427; 09JUK5; 38, Created)
 DT 15-JUL-1999 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Voltage-dependent T-type calcium channel alpha-1H subunit (Cav3.2).
 GN CACNA1H.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RA Miltman S.;
 RL "Exon organization of mouse Cacna1h.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1823-1952 FROM N.A.
 RC STRAIN-CS7BL/6J;
 RA Crittbs L.L., Lee J.-H., Yang J., Daud A.N., Perez-Reyes E.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDATE THE
 CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1H
 CC GIVES RISE TO T-TYPE CALCIUM CURRENTS. T-TYPE CALCIUM CHANNELS
 CC BELONG TO THE "LOW-VOLTAGE ACTIVATED (LVA)" GROUP AND ARE STRONGLY
 CC BLOCKED BY NICKEL AND MIBEFRADIL. A PARTICULARITY OF THIS TYPE OF
 CC CHANNELS IS AN OPENING AT OUTER NEGATIVE POTENTIALS AND A
 CC VOLTAGE-DEPENDENT INACTIVATION. T-TYPE CHANNELS SERVE PACEMAKING
 CC FUNCTIONS IN BOTH CENTRAL NEURONS AND CARDIAC NODAL CELLS AND
 CC SUPPORT CALCIUM SIGNALING IN SECRETORY CELLS AND VASCULAR SMOOTH
 CC MUSCLE. THEY MAY ALSO BE INVOLVED IN THE MODULATION OF FIRING
 CC PATTERNS OF NEURONS WHICH IS IMPORTANT FOR INFORMATION PROCESSING
 CC AS WELL AS IN CELL GROWTH PROCESSES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -1- PWM: IN RESPONSE TO RAISING OF INTRACELLULAR CALCIUM, THE T-TYPE
 CC CHANNELS ARE ACTIVATED BY CAM-KINASE II.
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF226868; AAK21607.2; -
 CC EMBL; AF026385; AAK21607.1; JOINED.
 CC EMBL; AF051947; AAC67240.1; -
 CC MGD; MGI:1928842; Cacna1h.
 CC DR InterPro: IPR001682; Ca/Na_pore.
 CC DR InterPro: IPR002111; Cal_channel_TTPl.
 CC DR InterPro: IPR005821; Ion_trans.
 CC DR InterPro: IPR005820; M-channel_nlg.
 CC DR Pfam: PF00520; Ion_trans. 4.
 CC DR Iodic channel: Transmembrane; Ion transport; Voltage-gated channel;
 CC Calcium channel; Glycoprotein; Repeat; Multigene family;
 CC Phosphorylation.
 CC KW REPEAT 87 422 I.
 CC REPEAT 87 422 II.
 CC REPEAT 776 1015 II.
 CC REPEAT 1292 1569 III.
 CC REPEAT 1613 1874 IV.
 CC DOMAIN 1 100 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 101 119 SI OF REPEAT I (POTENTIAL).
 CC

FT	DOMAIN	120	139	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	140	160	S2 OF REPEAT I (POTENTIAL).
FT	DOMAIN	161	169	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	170	184	S3 OF REPEAT I (POTENTIAL).
FT	DOMAIN	185	193	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	194	212	S4 OF REPEAT I (POTENTIAL).
FT	DOMAIN	213	232	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	233	253	S5 OF REPEAT I (POTENTIAL).
FT	DOMAIN	254	394	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	395	419	S6 OF REPEAT I (POTENTIAL).
FT	DOMAIN	420	790	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	791	811	S1 OF REPEAT II (POTENTIAL).
FT	DOMAIN	812	824	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	825	846	S2 OF REPEAT II (POTENTIAL).
FT	DOMAIN	847	852	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	853	871	S3 OF REPEAT II (POTENTIAL).
FT	DOMAIN	872	879	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	880	903	S4 OF REPEAT II (POTENTIAL).
FT	DOMAIN	904	914	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	915	935	S5 OF REPEAT II (POTENTIAL).
FT	DOMAIN	936	987	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	988	1012	S6 OF REPEAT II (POTENTIAL).
FT	DOMAIN	1013	1301	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1302	1324	S1 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1325	1342	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1343	1363	S2 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1364	1373	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1374	1393	S3 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1394	1407	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1408	1429	S4 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1430	1439	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1440	1463	S5 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1464	1540	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1541	1566	S6 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1567	1627	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1628	1648	S1 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1649	1662	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1663	1684	S2 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1685	1691	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1692	1710	S3 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1711	1724	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1725	1748	S4 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1749	1762	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1763	1783	S5 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1784	1846	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1847	1874	S6 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1875	2365	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	521	531	POLY-HIS.
FT	DOMAIN	1594	1597	POLY-ARG.
FT	SITE	378	378	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	971	971	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	1515	1515	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	SITE	1819	1819	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).
FT	CARBOHYD	192	192	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	271	271	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1477	1477	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CONFLICT	1823	1825	GIM -> ARG (IN REF. 2).
FT	CONFLICT	1914	1914	D -> E (IN REF. 2).
FT	CONFLICT	1945	1947	APA -> LLO (IN REF. 2).
FT	CONFLICT	1952	1952	S -> A (IN REF. 2).
FT	CONFLICT	1953	2351	MISSING (IN REF. 2).
SO	SEQUENCE	2365	AA; 261944	MM; 9A8A1570C210596 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 2365;
 Best Local Similarity 71.4%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WASTRES 7
 1:111

Db 1276 WCSSRES 1282

RESULT 14
POLG_TBEVH STANDARD; PRT: 3414 AA.

AC 001299;
ID POLG_TBEVH STANDARD; PRT: 3414 AA.

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Capsid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NS1, NS2a, NS2b, NS4a and NS4b; Protease/helicase (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5)].
DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
DE (NS5)].
OS tick-borne encephalitis virus (strain Hypr) (TBEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=70733;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallner G., Mandl C.W., Ecker M., Holzmann H., Stalsny K.,
RA Kunz C., Heinz F.X.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=91303656; PubMed=1712858;
RT Mandl C.W., Kunz C., Heinz F.X.;
RT *Presence of poly(A) in a flavivirus: significant differences between
RT the 3' noncoding regions of the genomic RNAs of tick-borne
RT encephalitis virus strains*;
RL J. Virol. 65:4070-4077(1991).
CC -1- FUNCTION: THE SMALL PROTEINS NS2a, NS2b, NS4a AND NS4b ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the p6
CC position, Cys or Thr in p1 and Ser or Ala in p1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
CC -1- MISCELLANEOUS: THE NONSTRUCTURAL PROTEINS NS1 PRESENTS TWO
CC ALTERNATIVE CLEAVAGE SITES FOR ITS C-TERMINUS, WHICH MAY DEFINE A
CC SOLUBLE OR A MEMBRANE-BOUND FORM OF NS1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39293; AAB53095.1; -
CC EMBL: M76660; AAA47904.1; -
CC HSSP: P14336; ISVB.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000336; Flavi_capsidC.
DR InterPro: IPR001122; Flavi_glycoprote.
DR InterPro: IPR001850; Flavi_helicase.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR001157; Flavi_NS1.
DR InterPro: IPR000752; Flavi_NS2a.
DR InterPro: IPR000487; Flavi_NS2b.
DR InterPro: IPR000404; Flavi_NS4a.
DR InterPro: IPR001528; Flavi_NS4b.
DR InterPro: IPR000208; Flavi_NS5.
DR InterPro: IPR002535; Flavi_propep.
DR InterPro: IPR002877; FtsJ.
DR InterPro: IPR001650; Helicase_C.

DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PsYfr.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF02832; Flavi_glycoprot; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF00949; Flavi_helicase; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF00948; Flavi_NS1; 1.
DR Pfam: PF01005; Flavi_NS2a; 1.
DR Pfam: PF01002; Flavi_NS2b; 1.
DR Pfam: PF01350; Flavi_NS4a; 1.
DR Pfam: PF01349; Flavi_NS4b; 1.
DR Pfam: PF00972; Flavi_NS5; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR Pfam: PF01728; FtsJ; 1.
DR Pfam: PF00271; helicase_C; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR ProDom: PD001496; Flavi_NS1; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC_C; 1.
DR PolyProtein: Glycoprotein; Transferase: RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein.
ET INIT_MER 1 1
ET CHAIN 1 112
ET PROPEP 113 205
ET CHAIN 206 280
ET CHAIN 281 776
ET CHAIN 777 2112
ET CHAIN 2112 1358
ET CHAIN 1359 1489
ET CHAIN 1490 2110
ET CHAIN 2111 2259
ET CHAIN 2260 2511
ET CHAIN 2512 3414
ET NP_BIND 1688 1695
ET SITE 1779 1782
ET TRANSMEM 101 112
ET TRANSMEM 247 259
ET TRANSMEM 266 280
ET TRANSMEM 738 751
ET DISULFID 283 310
ET DISULFID 340 396
ET DISULFID 354 385
ET DISULFID 372 401
ET DISULFID 466 570
ET DISULFID 587 618
ET CARBOHYD 144 144
ET CARBOHYD 434 434
ET CARBOHYD 861 861
ET CARBOHYD 983 983
ET CARBOHYD 999 999
ET CARBOHYD 1649 1649
ET CARBOHYD 1988 1988
ET CARBOHYD 2044 2044
ET CARBOHYD 2447 2447
ET CARBOHYD 2529 2529
ET CARBOHYD 2726 2726
SO SEQUENCE 3414 AA; 378539 MW; EC0B1A5325A08C19 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 3414;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 2931 WASARCA 2937

RESULT 15
POLG_TBEVH STANDARD; PRT: 3414 AA.

AC P14336; 088493; 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polypeptide [contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural
 DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
 DE (EC 3.4.21.98); RNA-directed RNA polymerase (EC 2.7.7.48)
 DE (NS5)].
 OS Tick-borne encephalitis virus (Western subtype) (TBEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_Taxid=11086;
 RN [1]
 RP SEQUENCE FROM N.A., AND REVISIONS.
 RC STRAIN=Neudoerfl; PubMed=7483260;
 RX MEDLINE=96036491; PubMed=7483260;
 RA Wallner G., Mandl C.W., Kunz C., Heinz F.X.;
 RT The flavivirus 3'-noncoding region: extensive size heterogeneity
 RT independent of evolutionary relationships among strains of tick-borne
 RT encephalitis virus.";
 RL Virology 213:169-178(1995).
 RN [2]
 RP SEQUENCE OF 1-779 FROM N.A.
 RC STRAIN=Neudoerfl; PubMed=3413985;
 RX MEDLINE=88322870; PubMed=3413985;
 RA Mandl C.W., Heinz F.X., Kunz C.;
 RT *Sequence of the structural proteins of tick-borne encephalitis virus
 RT (Western subtype) and comparative analysis with other flaviviruses.";
 RL Virology 166:197-205(1988).
 RN [3]
 RP SEQUENCE OF 767-3414 FROM N.A.
 RC STRAIN=Neudoerfl; PubMed=2554575;
 RX MEDLINE=90051080; PubMed=2554575;
 RA Mandl C.W., Heinz F.X., Stoeckl E., Kunz C.;
 RT *Genome sequence of tick-borne encephalitis virus (Western subtype)
 RT and comparative analysis of nonstructural proteins with other
 RT flaviviruses.";
 RL Virology 173:291-301(1989).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 281-680.
 RX MEDLINE=95272700; PubMed=753193; PubMed=753193;
 RA Rey F.A., Heinz F.X., Mandl C.W., Harrison S.C.;
 RT The envelope glycoprotein from tick-borne encephalitis virus at 2-A
 RT resolution.";
 RL Nature 375:291-298(1995).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polypeptide, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N). THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPID-PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U27495; AAA86870.1; -.
 CC DR PIR: A31052; GNMVNE.
 CC DR PDB: 1SVB; 10-JUN-96.
 CC DR MEROPS: S07; UPW; -.
 CC DR InterPro: IPR001410; DEAD.
 CC DR InterPro: IPR001122; Flav1_capsidC.

DR InterPro: IPR000336; Flav1_glycoprote.
 DR InterPro: IPR001850; Flav1_helicase.
 DR InterPro: IPR000069; Flav1_M.
 DR InterPro: IPR001157; Flav1_NS1.
 DR InterPro: IPR000752; Flav1_NS2A.
 DR InterPro: IPR000487; Flav1_NS2B.
 DR InterPro: IPR000404; Flav1_NS4A.
 DR InterPro: IPR001528; Flav1_NS4B.
 DR InterPro: IPR002028; Flav1_NS5.
 DR InterPro: IPR002535; Flav1_NS5.
 DR InterPro: IPR002877; FtsJ.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01003; Flav1_capsid; 1.
 DR Pfam: PF02832; Flav1_glycop_C; 1.
 DR Pfam: PF00869; Flav1_glycoprot; 1.
 DR Pfam: PF00949; Flav1_helicase; 1.
 DR Pfam: PF01004; Flav1_M; 1.
 DR Pfam: PF00948; Flav1_NS1; 1.
 DR Pfam: PF01005; Flav1_NS2A; 1.
 DR Pfam: PF01002; Flav1_NS2B; 1.
 DR Pfam: PF01350; Flav1_NS4A; 1.
 DR Pfam: PF01349; Flav1_NS4B; 1.
 DR Pfam: PF00972; Flav1_NS5; 1.
 DR Pfam: PF01570; Flav1_propep; 1.
 DR Pfam: PF01728; FtsJ; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR ProDom: PD001556; Flav1_glycoprote; 1.
 DR ProDom: PD001496; Flav1_NS1; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC_C; 1.
 KW Core protein; Coat protein; Envelope protein; RNA-directed RNA polymerase;
 KW App-binding; Transmembrane; Nonstructural protein; 3D-structure;
 KW Int_Met 1
 FT CHAIN 1
 FT PROPEP 113 205
 FT CHAIN 206 280
 FT CHAIN 281 776
 FT CHAIN 777 1128
 FT CHAIN 1129 1358
 FT CHAIN 1359 1489
 FT CHAIN 1490 2110
 FT CHAIN 2111 2259
 FT CHAIN 2260 2511
 FT CHAIN 2512 3414
 FT DOMAIN 378 391
 FT NP_BIND 1688 1695
 FT SITE 1779 1782
 FT TRANSMEM 101 112
 FT TRANSMEM 247 259
 FT TRANSMEM 266 280
 FT TRANSMEM 738 751
 FT DISULFID 283 310
 FT DISULFID 340 396
 FT DISULFID 354 385
 FT DISULFID 372 401
 FT DISULFID 466 570
 FT DISULFID 587 618
 FT CARBOHYD 144 144
 FT CARBOHYD 434 434
 FT CARBOHYD 861 861
 FT CARBOHYD 983 983
 FT CARBOHYD 999 999
 FT CARBOHYD 2447 2447
 FT HELIX 282 285
 FT STRAND 290 294
 FT STRAND 300 306
 FT TURN 307 308
 FT TURN 311 315
 FT TURN 316 317
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 ENVELOPE GLYCOPROTEIN M.
 MAJOR ENVELOPE PROTEIN E.
 NONSTRUCTURAL PROTEIN NS1.
 NONSTRUCTURAL PROTEIN NS2A.
 NONSTRUCTURAL PROTEIN NS2B.
 PROTEASE/HELICASE (NS3).
 NONSTRUCTURAL PROTEIN NS4A.
 NONSTRUCTURAL PROTEIN NS4B.
 RNA-DIRECTED RNA POLYMERASE (NS5).
 INVOLVED IN FUSION.
 DEAD BOX.
 HYDROPHOBIC SIGNAL SEQUENCE (POTENTIAL).
 POTENTIAL.
 POTENTIAL.
 POTENTIAL.

FT	STRAND	318	331
FT	STRAND	334	340
FT	STRAND	343	352
FT	TURN	355	356
FT	HELIIX	363	366
FT	TURN	368	369
FT	STRAND	370	379
FT	HELIIX	381	383
FT	TURN	384	384
FT	STRAND	389	401
FT	TURN	403	404
FT	STRAND	406	411
FT	TURN	414	416
FT	STRAND	418	425
FT	TURN	433	434
FT	STRAND	438	439
FT	STRAND	441	446
FT	TURN	447	448
FT	STRAND	449	449
FT	STRAND	451	455
FT	HELIIX	457	459
FT	STRAND	460	467
FT	HELIIX	468	470
FT	TURN	475	476
FT	STRAND	477	481
FT	TURN	484	485
FT	TURN	487	488
FT	STRAND	492	496
FT	HELIIX	497	501
FT	TURN	502	502
FT	STRAND	507	508
FT	TURN	510	511
FT	STRAND	516	516
FT	TURN	517	517
FT	HELIIX	518	521
FT	STRAND	522	524
FT	STRAND	528	528
FT	TURN	529	530
FT	STRAND	531	531
FT	STRAND	534	536
FT	HELIIX	541	547
FT	TURN	548	550
FT	STRAND	553	557
FT	TURN	558	559
FT	STRAND	560	562
FT	STRAND	567	573
FT	TURN	575	576
FT	STRAND	577	577
FT	TURN	582	583
FT	STRAND	586	586
FT	TURN	589	590
FT	STRAND	592	600
FT	STRAND	606	612

Query Match 78.9%; Score 30; DB 1; Length 3414;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WASTRES 7
||| ||:
Db 2931 WASAREA 2937

Search completed: October 7, 2003, 19:14:58
Job time : 2.42414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003, Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 6.97586 Seconds
(without alignments)
258.946 Million cell updates/sec

Title: us-09-988-013a-2_COPY_56_62
Perfect score: 38
Sequence: 1 WASTRES 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_minc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	92.1	130	4 Q9NP29	Q9NP29 homo sapien
2	34	89.5	980	16 Q9PL66	Q9PL66 xanthomonas
3	33	86.8	1317	3 Q8TF23	Q8TF23 aspergillus
4	32	84.2	347	16 Q915J7	Q915J7 pseudomonas
5	31	81.6	171	5 Q9NSV2	Q9NSV2 caenorhabdi
6	31	81.6	183	2 Q57000	Q57000 zymomonas m
7	31	81.6	226	2 Q88087	Q88087 enterococu
8	31	81.6	230	16 Q8NU75	Q8NU75 corynebacte
9	31	81.6	404	2 Q9AMJ2	Q9AMJ2 streptomyce
10	31	81.6	564	5 Q9NA93	Q9NA93 caenorhabdi
11	31	81.6	572	16 Q92X10	Q92X10 rhizobium m
12	31	81.6	642	2 Q8RMG0	Q8RMG0 nostoc sp.
13	31	81.6	642	16 Q9A805	Q9A805 caulobacter
14	31	81.6	642	16 Q8Z0C9	Q8Z0C9 anabaena sp
15	31	81.6	727	5 Q9BNW8	Q9BNW8 scutigerell
16	31	81.6	735	16 Q989J7	Q989J7 rhizobium 1

17	31	81.6	783	2 Q8KJN0	Q8KJN0 rhizobium 1
18	31	81.6	1499	10 Q9LH44	Q9LH44 arabidopsis
19	30	78.9	29	11 Q8C837	Q8C837 mus musculu
20	30	78.9	51	4 Q9UJ11	Q9UJ11 homo sapien
21	30	78.9	62	12 Q8B8F5	Q8B8F5 hepatitis b
22	30	78.9	103	17 Q8Z272	Q8Z272 pyrobaculum
23	30	78.9	118	6 Q9BGP0	Q9BGP0 macaca fasc
24	30	78.9	119	12 Q8JYX1	Q8JYX1 hepatitis b
25	30	78.9	119	12 Q8JYX2	Q8JYX2 hepatitis b
26	30	78.9	119	12 Q8JYX3	Q8JYX3 hepatitis b
27	30	78.9	130	4 Q8MWL9	Q8MWL9 homo sapien
28	30	78.9	206	11 Q9CTW1	Q9CTW1 mus musculu
29	30	78.9	207	15 Q9UB47	Q9UB47 human immun
30	30	78.9	209	2 P94650	P94650 chlorobium
31	30	78.9	256	16 Q913X8	Q913X8 pseudomonas
32	30	78.9	285	16 Q9B838	Q9B838 rhizobium 1
33	30	78.9	301	12 Q8JMA5	Q8JMA5 manestra co
34	30	78.9	310	10 Q9C939	Q9C939 arabidopsis
35	30	78.9	349	4 Q8TC24	Q8TC24 homo sapien
36	30	78.9	349	11 Q9W001	Q9W001 mus musculu
37	30	78.9	356	10 Q8RYX9	Q8RYX9 oryza sativ
38	30	78.9	409	16 Q8UG04	Q8UG04 agrobacteri
39	30	78.9	415	17 Q8TFM2	Q8TFM2 metanosc
40	30	78.9	462	16 Q984H6	Q984H6 rhizobium 1
41	30	78.9	489	16 Q8UE58	Q8UE58 agrobacteri
42	30	78.9	528	5 Q8SVE2	Q8SVE2 encaphalito
43	30	78.9	558	17 Q8TQ86	Q8TQ86 methanosarc
44	30	78.9	650	16 Q8Y005	Q8Y005 ralsstonia s
45	30	78.9	677	11 Q8BP91	Q8BP91 mus musculu

ALIGNMENTS

RESULT 1

Q9NP29 ID Q9NP29 PRELIMINARY: PRT: 130 AA.

AC Q9NP29;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Microfibrillar protein 2 (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96354815; PubMed=8753791;

RA Ozsvath K.J., Xia S., Hirose H., Tilson M.D.;

RT "Two hypothetical proteins of human aortic adventitia, with Ig kappa, collagenous, and aromatic-rich motifs.";

RL Biochem. Biophys. Res. Commun. 225:500-504(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=97367690; PubMed=9224393;

RA Ozsvath K.J., Hirose H., Xia S., Chew D., Knoetgen J. III,

RT Tilson M.D.;

RT "Expression of two novel recombinant proteins from aortic adventitia (kappas) sharing amino acid sequences with cytomegalovirus.";

RN [3]

RP SEQUENCE FROM N.A.

RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF206020; AAF62402.1; -.

DR HSSP: P80362; IWTL.

FT NON_TER 130

SO SEQUENCE 130 AA; 14128 MW; 51275185ACGFAIE CRC64;

Query Match 92.1%; Score 35; DB 4; Length 130;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 WASTRES 7
       1111111
Db      76 WASTRDS 82

RESULT 2
Q8PL66 PRELIMINARY; PRT: 980 AA.
ID Q8PL66
AC Q8PL66;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE GGDFF family protein.
GN XAC1940.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;

RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-306 / ATCC 13902 / XV 101;
RX MEDLINE-22022145; PubMed-12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Clapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locelli E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Melanda J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sana J.A.D., Silva C., de Souza R.F.,
RA Spicola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.,
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE011829; AAM36802.1; -.
DR InterPro: IPR00160; GGDFF.
DR Pfam: PF00990; GGDFF. 1.
DR TIGRFAMs; TIGR00254; GGDFF. 1.
KM Complete proteome.
SQ SEQUENCE 980 AA; 107940 MW; 47179C81E6DCE442 CRC64;

Query Match
Best Local Similarity 89.5%; Score 34; DB 16; Length 980;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WASTRE 6
       1111111
Db      492 WASTRE 497

RESULT 3
Q8TF23 PRELIMINARY; PRT: 1317 AA.
ID Q8TF23
AC Q8TF23;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Conserved hypothetical protein.
DE AFA35310.09C.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillius.
OX NCBI_TaxID=5085;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Af293;

```

```

RA Knowles D.G., Warren T., Hall N., Quail M., Woodward J.R.,
RA Denning D.W., Anderson M.J., Bartell B.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL731691; CAD29602.1; -.
DR InterPro: IPR01680; WD40.
DR Pfam: PF00400; WD40; 2.
DR SMART; SM00320; WD40; 3.
KM Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 1317 AA; 148645 MW; 0DC4B0BD149A45DE CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 3; Length 1317;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WASTRES 7
       1111111
Db      979 WASTRES 985

RESULT 4
Q9ISJ7 PRELIMINARY; PRT: 347 AA.
ID Q9ISJ7
AC Q9ISJ7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein PA0732.
GN PA0732.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;

RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
RA Garner R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lapidis K., Llin R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004508; AAG04121.1; -.
DR InterPro: IPR005490; ERFK_YBIS_Yhng.
DR Pfam: PF03734; ERFK_YBIS_Yhng. 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 347 AA; 37161 MW; 264E4DC064290CC3 CRC64;

Query Match
Best Local Similarity 84.2%; Score 32; DB 16; Length 347;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 WASTRES 7
       1111111
Db      318 WASTRDT 324

RESULT 5
Q9NSV2 PRELIMINARY; PRT: 171 AA.
ID Q9NSV2
AC Q9NSV2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 17.6 kDa protein.
GN CC8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode *C. elegans*: a platform for
investigating biology. The *C. elegans* Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX Kemp K., Le T.T.;
RT "The sequence of *C. elegans* cosmid CC8";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX Waterston R.;
RT "Direct Submission";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006626; AAF39788.1;
DR WormPep; CC8.1; CE20616.
SO Hypothetical protein.
SQ SEQUENCE 171 AA; 17643 MW; 03AD996B9A627CFD CRC64;

Query Match 81.6%; Score 31; DB 5; Length 171;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6
DB 122 WASTRE 127

RESULT 6
ID 057000 PRELIMINARY; PRT; 183 AA.
AC 057000;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)
DE Secretion activator protein.
GN ZLITS.
OS *Zymomonas mobilis*.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; *Zymomonas*.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IP013756;
RX Kondo Y., Toyoda A., Fukushi H., Yanase H., Tonomura K., Kawasaki H.,
RA Sakai T.;
RT "Cloning and characterization of a pair of genes that stimulate the
production and secretion of *Zymomonas mobilis* extracellular
levansucrase and invertase.";
RL Biosci. Biotechnol. Biochem. 0:0-0(1994).
RL EMBL; D17522; BAA04473.1;
SQ SEQUENCE 183 AA; 20629 MW; F8F00828108BD66F CRC64;

Query Match 81.6%; Score 31; DB 2; Length 183;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 145 WASTREN 151

RESULT 7
ID 088087 PRELIMINARY; PRT; 226 AA.
AC 088087;
DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE ABC transporter ATP-binding protein.
OS *Enterococcus faecalis* (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; *Enterococcus*.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OG1RF;
RX MEDLINE=98380380; PubMed=9712783;
RA Xu Y., Murray B.E., Weinstock G.M.;
RT "A cluster of genes involved in polysaccharide biosynthesis from
Enterococcus faecalis OG1RF";
RL Infect. Immun. 66:4313-4323(1998).
DR EMBL; AF071085; AAC35926.1;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_Transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_Transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding.
SQ SEQUENCE 226 AA; 25420 MW; 68C6D5CE85CD9A2A5 CRC64;

Query Match 81.6%; Score 31; DB 2; Length 226;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
DB 219 WASTRES 225

RESULT 8
ID 08NU75 PRELIMINARY; PRT; 230 AA.
AC 08NU75;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Nicotinamide mononucleotide transporter.
GN GGL0064.
OS *Corynebacterium glutamicum* (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; *Corynebacterium*.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005274; BAB97457.1;
DR InterPro; IPR007056; NNN_Transporter.
DR InterPro; IPR006419; NNN_trans_Pnuc.
DR Pfam; PF04973; NNN_Transporter; 1.
DR TIGRFAMS; TIGR01528; NNN_trans_Pnuc; 1.
KW Complete proteome.
SQ SEQUENCE 230 AA; 25546 MW; 93429B9438A8AE45 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 230;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6
DB 113 WASTRE 118

RESULT 9
ID 09AMJ2 PRELIMINARY; PRT; 404 AA.
AC 09AMJ2;
DT 09AMJ2;

DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE Putative ketosynthase Sima2 (Sim3).
 GN SIMA2 OR SIM3.
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 RX NCBI_TaxID=1890;
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tue6040;
 RA Trefzer A., Bechthold A.;
 RT "The simocyclinone biosynthetic gene cluster isolated from
 RT Streptomyces antibioticus Tue6040."
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RM [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tu 6040;
 RA Gain U., Li S.-M., Schimana J., Fiedler H.-P., Helde L.;
 RT "Identification of the Simocyclinone Biosynthetic Gene Cluster of
 RT Streptomyces antibioticus Tu 6040."
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF324838; AAK06785.1; -
 DR EMBL; AF322256; AAL15581.1; -
 DR HSP; P73283; 1ESM.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR Pfam: PF00109; ketoacyl-synt_1.
 DR Pfam: PF02801; ketoacyl-synt_C_1.
 SQ SEQUENCE 404 AA; 41905 MW; CDE4A9DD54DD6B8 CRC64;

Query Match 81.6%; Score 31; DB 2; Length 404;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6
 |||||
 Db 24 WASTRE 29

RESULT 10
 ID Q9NA93 PRELIMINARY; PRT; 564 AA.
 AC Q9NA93;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Y53F4B.27 protein.
 GN Y53F4B.27
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RX NCBI_TaxID=6239;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Smye R.;
 RC Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AL132949; CAB61101.2; -
 DR WormPep: Y53F4B.27; CE26159.
 SQ SEQUENCE 564 AA; 62463 MW; 950BB275855EF1C8 CRC64;

Query Match 81.6%; Score 31; DB 5; Length 564;
 Best Local Similarity 83.3%; Pred. No. 3.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6

Db 33 WASTRE 38

RESULT 11
 ID Q92X10 PRELIMINARY; PRT; 572 AA.
 AC Q92X10;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Hypothetical protein RB0152.
 GN RB0152 OR SMD20152.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 RX NCBI_TaxID=382;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE-21396508; PubMed-11481431;
 RA Finn T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
 RA Golding B., Puenler A.;
 RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL063642; CAC48552.1; -
 DR InterPro: IPR004843; M-peptidase.
 DR Pfam: PF00149; Metallophos; 1.
 DR PfamId: Hypothetical protein; Complete proteome.
 SQ SEQUENCE 572 AA; 62894 MW; B0E25EC10F213353 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 572;
 Best Local Similarity 71.4%; Pred. No. 3.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
 |||||
 Db 41 WASTROS 47

RESULT 12
 ID Q8RMG0 PRELIMINARY; PRT; 642 AA.
 AC Q8RMG0;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Cyclomaltoextrin glucanotransferase.
 OS Nostoc sp. (strain PCC 9229).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 RX NCBI_TaxID=70817;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Wouters J., Janson S., Bergman B.;
 RC STRAIN-PCC 9229;
 RT "Cloning and of expression of a putative cyclomaltoextrin
 RT glucanotransferase from the symbiotically competent cyanobacterium
 RT Nostoc sp. PCC 9229."
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF497477; AAM16154.1; -
 DR InterPro: IPR006048; Alpha-amy1-C.
 DR InterPro: IPR006047; Alpha-amy1-C.
 DR InterPro: IPR006589; Alp-amy1-cat.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amy1ase; 1.
 DR Pfam: PF02806; alpha-amy1ase_C; 1.
 DR Pfam: PF00686; CBM_20; 1.
 DR SMART: SM00642; Amy; 1.
 KW Transglutase.
 SQ SEQUENCE 642 AA; 74063 MW; A3806A00720797B6 CRC64;

Query Match 81.6%; Score 31; DB 2; Length 642;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
 DB 636 WASTRES 642

RESULT 13

O9A805 PRELIMINARY; PRT; 642 AA.
 AC O9A805;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE IOLC protein.
 GN CCL1298.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblum T.V., Laub M.R., Ohta N., Maddock J.R.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AEO05805; AAK23279.1; -.
 DR TIGR: CCL298; -.
 DR InterPro: IPR002173; PFKB.
 DR Pfam: PF00294; PFKB; 1.
 DR PROSITE: PS00583; PFKB_KINASES_1; 1.
 DR PROSITE: PS00584; PFKB_KINASES_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 642 AA; 70888 MW; 2410443215415889 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 642;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WASTRE 6
 DB 352 WASTRE 357

RESULT 14

O8Z0C9 PRELIMINARY; PRT; 642 AA.
 AC O8Z0C9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Cyclomalto-dextrin glucanotransferase.
 GN ALR0169.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Matanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsuno M., Matsuno M., Muraiki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,

RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003581; BAB77693.1; -.
 DR InterPro: IPR006048; Alpha_amy1_C.
 DR InterPro: IPR006047; Alpha_amy1_cat.
 DR InterPro: IPR006589; Alp_amy1_cat_sub.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR006046; Glyco_hydro_13.
 DR Pfam: PF00128; alpha-amy1ase; 1.
 DR Pfam: PF02806; alpha-amy1ase_C; 1.
 DR Pfam: PF00686; CBM_20; 1.
 DR PRINTS: PR00110; ALPHAMYLASE.
 DR SMART: SM00642; Amy; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 642 AA; 74287 MW; 9F0991C0452F9F1C CRC64;

Query Match 81.6%; Score 31; DB 16; Length 642;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
 DB 636 WASTRES 642

RESULT 15

O9BNW8 PRELIMINARY; PRT; 727 AA.
 AC O9BNW8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Elongation factor-2 (Fragment).
 GN Scutigere1la sp. Scuz2.
 OS Scutigere1la sp. Scuz2.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Symphyla; Scutigerellidae;
 OC Scutigerella.
 OX NCBI_TaxID=109756;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21317060; PubMed=11421654;
 RA Regier J.C., Shultz J.W.;
 RT "Elongation factor-2: a useful gene for arthropod phylogenetics.";
 RL Mol. Phylogenet. Evol. 20:136-148(2001).
 DR EMBL: AF240827; AAK12352.1; -.
 DR HSSP: P13551; IDAR.
 DR InterPro: IPR005517; EFG_IV.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR000795; EF_GTPbind.
 DR Pfam: PF03764; EFG_IV; 1.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF03144; GTP_EFTU_D2; 1.
 DR PRINTS: PR00315; ELONGACTRECT.
 DR PROSITE: PS00301; EFATOR_GTP; 1.
 KW GTP-binding; Protein biosynthesis.
 FT NON_TER 1
 FT NON_TER 727
 SQ SEQUENCE 727 AA; 81422 MW; 577371F1AA122178 CRC64;

Query Match 81.6%; Score 31; DB 5; Length 727;
 Best Local Similarity 71.4%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WASTRES 7
 DB 256 WASTRES 262

Search completed: October 7, 2003, 19:19:57
 Job time : 8.97586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:26 ; Search time 10.069 Seconds
(without alignments)
126.112 Million cell updates/sec

Title: US-09-988-013a-2_COPY_95_102
Perfect score: 48
Sequence: 1 HQYLSSWT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	112	15	AA54932
2	48	100.0	112	15	AA54933
3	48	100.0	113	17	AA54932
4	48	100.0	113	17	AA54932
5	48	100.0	113	18	AA54932
6	48	100.0	115	18	AA54932
7	48	100.0	263	20	AA54932
8	48	100.0	268	20	AA54932
9	48	100.0	268	20	AA54932

10	48	100.0	273	20	AA54932	Anti-B7.1/anti-B7.1
11	48	100.0	301	18	AA54932	Single chain, huma
12	48	100.0	301	20	AA54932	Multispecific sing
13	48	100.0	301	22	AA54932	Single chain huma
14	48	100.0	301	22	AA54932	Single chain huma
15	48	100.0	352	20	AA54932	Anti-Fc gamma rece
16	48	100.0	553	18	AA54932	Single chain anti-
17	48	100.0	553	20	AA54932	H22-anti-CPA anti-
18	48	100.0	553	22	AA54932	Bispecific single
19	48	100.0	553	22	AA54932	Bispecific single
20	48	100.0	556	20	AA54932	Bispecific tetrava
21	48	100.0	580	20	AA54932	Bispecific tetrava
22	48	100.0	580	20	AA54932	Synthetic mouse fu
23	48	100.0	86	22	AA54932	Human kidney relat
24	48	100.0	86	22	AA54932	Anti-ClfA monoclon
25	48	100.0	86	22	AA54932	S. aureus ClfA spe
26	48	100.0	112	23	AA54932	S. aureus ClfA spe
27	48	100.0	112	23	AA54932	S. aureus ClfA spe
28	48	100.0	112	23	AA54932	Pinus radiata cell
29	48	100.0	154	21	AA54932	Pinus radiata cell
30	48	100.0	154	21	AA54932	Mouse apoptosis-in
31	48	100.0	154	21	AA54932	Mouse apoptosis-in
32	48	100.0	154	21	AA54932	Human male enhance
33	48	100.0	154	21	AA54932	A. thaliana enviro
34	48	100.0	154	21	AA54932	Eucahyptus grandis
35	48	100.0	154	21	AA54932	Staphylococcus aur
36	48	100.0	154	21	AA54932	Staphylococcus aur
37	48	100.0	154	21	AA54932	Novel human diagn
38	48	100.0	154	21	AA54932	CCR3-binding phage
39	48	100.0	154	21	AA54932	Human anti-placent
40	48	100.0	154	21	AA54932	S. aureus ClfA spe
41	48	100.0	154	21	AA54932	Human polypeptide
42	48	100.0	154	21	AA54932	Drosophila melanog
43	48	100.0	154	21	AA54932	Candida albicans e
44	48	100.0	154	21	AA54932	
45	48	100.0	154	21	AA54932	

ALIGNMENTS

RESULT 1
ID AAR54932 standard; peptide: 112 AA.
AC AAR54932:
XX
DT 25-MAR-2003 (updated)
DT 19-OCT-1994 (first entry)
XX
DE Fc receptor humanized VK chain 022 HuVK.
XX
KW Fc receptor; FcR; humanized antibody; hAb; IgG; cancer; allergy;
KW autoimmune disease; heteroantibody; bifunctional antibody;
KW immunotoxin; CDR; complementarity determining region; VH;
KW heavy chain variable region; VK; kappa chain variable region;
KW monoclonal phagocyte; PCR; polymerase chain reaction; primer;
KW site-directed mutagenesis; HuVK; monoclonal antibody; hAb.
XX
OS Homo sapiens; Mus sp.
XX
PN WO9410332-A1.
XX
PD 11-MAY-1994.
XX
PF 04-NOV-1993; 93MO-US10384.
XX
PR 04-NOV-1992; 92GB-0023377.
XX
PA (MEDA-) MEDAREX INC.
XX
PI Carr FU, Harris WJ, Tempest PR.
XX

CC recombinantly linked to the framework sequences of human VK and VH regions, respectively, to give humanised LL2 VK (AAR92217) and VH (AAR92218). These were subsequently linked, respectively, to human kappa and IgG1 constant regions. A humanised Mab was obtd. that retained the B-lymphoma and leukaemia cell targeting and CC internalisation characteristics of the parental LL2 Mab, and which exhibited a lowered HAMA reaction. It can be linked to e.g. a cytostatic agent for therapeutic appln.

CC Sequence 113 AA:

Query Match 100.0%; Score 48; DB 17; Length 113;

Best Local Similarity 100.0%; Pred. No. 0.59; Mismatches 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
| | | | | | | |

DB 95 HOYLSSWT 102

RESULT 4

AAR92217 standard; Protein: 113 AA.

AC AAR92217;

DT 28-MAY-1996 (first entry)

XX Humanised LL2 Mab VK region.

XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma; leukaemia; therapy; diagnosis; complementarity determining region;

KW CDR; antibody engineering.

XX Synthetic.

FT Key Location/Qualifiers

FT Region 24..40 /label= CDR1

FT Region 56..62 /label= CDR2

FT Region 95..103 /label= CDR3

XX MO9604925-A1.

PD 22-FEB-1996.

XX 11-AUG-1995; 95WO-US09641.

XX 12-AUG-1994; 94US-0289576.

XX (IMMU-) IMMUNOMEDICS INC.

PI Hansen H, Leung S;

DR WPI; 1996-139454/14.

XX N-PSDB; AAT15803.

PT Chimeric and humanised LL2 antibodies - used to produce conjugates for the therapy and diagnosis of B-cell lymphoma(s) and

PT Leukaemia(s).

XX Claim 5; Page 38; 70pp; English.

PS The complementarity determining regions (CDRs) of mouse monoclonal antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were recombinantly linked to the framework sequences of human VK and VH regions, respectively, to give humanised LL2 VK (AAR92217) and VH (AAR92218). These were subsequently linked, respectively, to human kappa and IgG1 constant regions. A humanised Mab was obtd. that retained the B-lymphoma and leukaemia cell targeting and internalisation characteristics of the parental LL2 Mab, and which

CC exhibited a lowered HAMA reaction. It can be linked to e.g. a cytostatic agent for therapeutic appln.

CC Sequence 113 AA:

Query Match 100.0%; Score 48; DB 17; Length 113;

Best Local Similarity 100.0%; Pred. No. 0.59; Mismatches 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
| | | | | | | |

DB 95 HOYLSSWT 102

RESULT 5

AAW27695 standard; Protein: 113 AA.

XX AAW27695;

DT 14-APR-1998 (first entry)

XX Variable kappa chain of Mab LL2.

XX Variable kappa chain; B cell; monoclonal antibody; Mab; LL2; B cell lymphoma; lymphocytic leukaemia cell; murine;

KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma; chronic lymphocytic leukaemia.

XX Mus sp.

OS Synthetic.

FT Key Location/Qualifiers

FT Region 18..20 /note= "potential N-linked glycosylation site"

FT Region 24..40 /note= "complementarity determining region 1"

FT Region 56..62 /note= "complementarity determining region 2"

FT Region 95..102 /note= "complementarity determining region 3"

XX MO9734632-A1.

PD 25-SEP-1997.

XX 19-MAR-1997; 97WO-US04196.

XX 20-MAR-1996; 96US-0013709.

XX (IMMU-) IMMUNOMEDICS INC.

PI Hansen H, Leung S, Qu Z;

DR WPI; 1997-479995/44.

XX N-PSDB; AAT88128.

PT Monoclonal antibody engineered to contain glycosylation site - in non-Fc constant heavy or light chain region, useful to diagnose or

PT treat B cell malignancies, e.g. non-Hodgkins lymphoma

XX Example 3; Fig 4A; 88pp; English.

PS The present sequence is the variable kappa chain of the B cell specific monoclonal antibody (Mab) LL2, which contains an engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell murine Mab. The Mab can be used to diagnose or treat B cell malignancies, e.g. non-Hodgkins lymphoma or chronic lymphocytic leukaemia. The glycosylation site allows a label or therapeutic agent of increased size to be conjugated to the carbohydrate moiety, without affecting the Mab's binding affinity or specificity.

```

XX      SQ      Sequence      113 AA;
Query Match
Best Local Similarity      100.0%; Score 48; DB 18; Length 113;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 HOYLSSWT 8
      11111111
      95 HOYLSSWT 102

RESULT 6
ID      AAW27697 standard; Protein: 115 AA.
XX      AAW27697;
XX      14-APR-1998 (first entry)
XX      Variable kappa chain of MAb hLL2.
XX      Variable kappa chain; B cell; monoclonal antibody; MAb: hLL2;
XX      B cell lymphoma; lymphocytic leukaemia cell; murine; humanised;
XX      diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
XX      chronic lymphocytic leukaemia.
XX      Chimeric - Mus sp.
XX      Chimeric - Homo sapiens.
XX      Key
XX      Region      24..40      Location/Qualifiers
XX      Region      56..62      /note= "complementarity determining region 1"
XX      Region      95..102     /note= "complementarity determining region 2"
XX      Region      /note= "complementarity determining region 3"

PN      WO9734632-A1.
XX      25-SEP-1997.
XX      19-MAR-1997;      97WO-US04196.
XX      20-MAR-1996;      96US-0013709.
XX      (IMMU-) IMMUNOMEDICS INC.
XX      Hansen H, Leung S, Qu Z;
XX      WPI: 1997-479995/44.
XX      N-PSDB; AAT8130.
XX      Monoclonal antibody engineered to contain glycosylation site - in
XX      non-Fc constant heavy or light chain region, useful to diagnose or
XX      treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX      Example 3; Fig 5a; 88pp; English.
XX      The present sequence is the variable kappa chain of the
XX      B cell specific monoclonal antibody (MAb) hLL2. hLL2 is a highly
XX      specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
XX      humanised murine MAb. The MAb can be used to diagnose or treat B
XX      cell malignancies, e.g. non-Hodgkins lymphoma or chronic
XX      lymphocytic leukaemia.
XX      Sequence      115 AA;

Query Match
Best Local Similarity      100.0%; Score 48; DB 18; Length 115;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 HOYLSSWT 8

```

```

DB      95 HOYLSSWT 102
      11111111
      95 HOYLSSWT 102

RESULT 7
ID      AAW90226 standard; Protein: 263 AA.
XX      AAW90226;
XX      10-MAY-1999 (first entry)
XX      Anti-B7.2 monospecific triabody 1G10.
XX      B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
XX      CD86; T cell activation; inhibitor; graft versus host disease;
XX      transplant rejection; allograft rejection; autoimmune disease;
XX      allergy; therapy; human; triabody; antibody; 1G10.
XX      Chimeric - Mus sp.
XX      Chimeric - synthetic.
XX      Key
XX      Peptide      1..24      Location/Qualifiers
XX      Region      25..144     /note= "peab signal peptide"
XX      Peptide      145..257   /note= "anti B7.2 MAb VH region"
XX      Peptide      258..263   /note= "anti B7.2 MAb VL region"
XX      Peptide      /note= "His6 tag"

PN      WO9858965-A2.
XX      30-DEC-1998.
XX      22-JUN-1998;      98WO-EP03791.
XX      20-JUN-1997;      97EP-0870092.
XX      (INNO-) INNOGENETICS NV.
XX      Bosman A, Buyse M, Lotte K, Sablon E;
XX      WPI: 1999-105615/09.
XX      N-PSDB; AAX01660.
XX      New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX      immune diseases including allograft rejection
XX      Example 7.3; Fig 34; 182pp; English.
XX      This polypeptide comprises a 1G10 monospecific triabody composed
XX      of the VH region of anti-B7.2 monoclonal antibody (MAb) 1G10
XX      joined to the VL region of 1G10. A triabody is a mono- a bi- or
XX      a trispecific molecule recognising simultaneously e.g. two B7.2
XX      and one B7.1 molecules. It has a rigid structure that prevents
XX      simultaneous binding to the 3 targets. Each antigen-binding site
XX      is formed by pairing of one VH and one VL domain from the same or
XX      from two different polypeptides. The invention relates to novel
XX      molecules, including triabodies, which can cross-link and/or
XX      cross-react with the costimulatory molecules B7.1 and B7.2 expressed
XX      on professional antigen-presenting cells, leading to the inhibition
XX      of antigen-specific T cell activation. Methods are provided for
XX      the production of such B7-binding molecules, and for their use in
XX      the treatment or prevention of diseases of the immune system, in
XX      particular graft rejection, graft versus host disease, allergy and
XX      autoimmune diseases (claimed).
XX      Sequence      263 AA;

Query Match
Best Local Similarity      100.0%; Score 48; DB 20; Length 263;

```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HOYLSSWT 8
 |||||
 Db 239 HOYLSSWT 246

RESULT 8
 AAW90228
 ID AAW90228 standard; Protein; 268 AA.

AAW90228;

10-MAY-1999 (first entry)

Anti-B7.1/anti-B7.2 bispecific triabody II.

B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 CD86; T cell activation; inhibitor; graft versus host disease;
 transplant rejection; allograft rejection; autoimmune disease;
 allergy; therapy; human; triabody; antibody; B7-24; 1G10.

Chimeric - Mus sp.
 Chimeric - synthetic.

Key Location/Qualifiers

Peptide 1..39
 /note="g3p signal peptide"

Region 40..155
 /note="anti B7.1 MAb VH region"

Peptide 156..268
 /note="anti B7.2 MAB VL region"

WO9858965-A2.

30-DEC-1998.

22-JUN-1998; 98WO-EP03791.

20-JUN-1997; 97EP-0870092.

(INNO-) INNOGENETICS NV.

Bosman A, Buyse M, Lorre K, Sablon E;

WPI: 1999-105615/09.
 N-PSDB; AAX01662.

New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 immune diseases including allograft rejection

Example 7.3; Fig 38; 182pp; English.

This polypeptide comprises a bispecific triabody composed of the VH
 region of anti-B7.1 monoclonal antibody (Mab) B7-24 joined to the
 VL region of anti-B7.2 Mab 1G10. A triabody is a mono- a bi- or
 a trispecific molecule recognising simultaneously e.g. two B7.2
 and one B7.1 molecules. It has a rigid structure that prevents
 simultaneous binding to the 3 targets. Each antigen-binding site
 is formed by pairing of one VH and one VL domain from the same or
 from two different polypeptides. The invention relates to novel
 molecules, including triabodies, which can cross-link and/or
 cross-react with the costimulatory molecules B7.1 and B7.2 expressed
 on professional antigen-presenting cells, leading to the inhibition
 of antigen-specific T cell activation. Methods are provided for
 the production of such B7-binding molecules, and for their use in
 the treatment or prevention of diseases of the immune system, in
 particular graft rejection, graft versus host disease, allergy and
 autoimmune diseases (claimed).

Sequence 268 AA;

Query Match 100.0%; Score 48; DB 20; Length 268;

Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HOYLSSWT 8
 |||||
 Db 250 HOYLSSWT 257

RESULT 9
 AAW90222

ID AAW90222 standard; Protein; 268 AA.

AAW90222;

10-MAY-1999 (first entry)

Anti-B7.2 monospecific diabody 1G-10.

B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 CD86; T cell activation; inhibitor; graft versus host disease;
 transplant rejection; allograft rejection; autoimmune disease;
 allergy; therapy; human; diabody; antibody; 1G-10.

Chimeric - Mus sp.
 Chimeric - synthetic.

Key Location/Qualifiers

Peptide 1..24
 /note="pe1b signal peptide"

Region 25..144
 /note="anti B7.2 MAB VH region"

Peptide 145..149
 /note="G4S flexible linker"

Region 150..262
 /note="anti B7.2 MAB VL region"

Peptide 263..268
 /note="His6 tag"

WO9858965-A2.

30-DEC-1998.

22-JUN-1998; 98WO-EP03791.

20-JUN-1997; 97EP-0870092.

(INNO-) INNOGENETICS NV.

Bosman A, Buyse M, Lorre K, Sablon E;

WPI: 1999-105615/09.
 N-PSDB; AAX01656.

New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 immune diseases including allograft rejection

Example 7.2; Fig 26; 182pp; English.

This polypeptide comprises a 1G-10 monospecific diabody composed
 of the VH region of anti-B7.2 monoclonal antibody (Mab) 1G-10
 joined via a short, flexible linker to the VL region of 1G-10.
 Mono- or bispecific bivalent molecules are generated by shortening
 the flexible linker sequence between the VH and VL of the anti-B7.1
 scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with
 dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
 for bispecific molecules by cross-pairing the VH and VL domains
 from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
 B7.12/B7.12). The invention relates to novel molecules, including
 diabodies, which can cross-link and/or cross-react with the
 costimulatory molecules B7.1 and B7.2 expressed on professional
 antigen-presenting cells, leading to the inhibition of antigen-
 specific T cell activation. Methods are provided for the
 production of such B7-binding molecules, and for their use in the

CC treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).
 CC
 XX Sequence 268 AA;
 SQ

Query Match 100.0%; Score 48; DB 20; Length 268;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
 DB 244 HOYLSSWT 251

RESULT 10
 ID AAW90224 standard; Protein: 273 AA.
 XX AAW90224;
 AC AAW90224;
 XX 10-MAY-1999 (first entry)
 DT
 XX
 DE Anti-B7.1/anti-B7.2 bispecific diabody II.
 XX
 XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KW CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; diabody; antibody; B7-24; Ig10.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..39
 FT /note="g3p signal peptide"
 FT 40..155
 FT /note="anti B7.1 MAb VH region"
 FT 156..160
 FT /note="G4S flexible linker"
 FT 161..273
 FT /note="anti B7.2 MAB VL region"
 FT Misc-difference 21
 FT /note="encoded by TCA"
 FT
 PN WO9858965-A2.
 PD 30-DEC-1998.
 XX
 PD 22-JUN-1998; 98WO-EP03791.
 XX
 PR 20-JUN-1997; 97EP-0870092.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Bosman A, Buyse M, Lorre K, Sablon E;
 DR WPI: 1999-105615/09.
 DR N-PSDB; AAX01658.
 XX
 PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 XX
 XX Example 7.2; Fig 30; 182pp; English.
 PS
 CC This polypeptide comprises a bispecific diabody composed of the VH
 CC region of anti-B7.1 monoclonal antibody (MAb) B7-24 joined via a
 CC short, flexible linker to the VL region of anti-B7.2 MAb Ig10.
 CC Mono- or bispecific bivalent molecules are generated by shortening
 CC the flexible linker sequence between the VH and VL of the anti-B7.1
 CC scfv B7-24, the anti-B7.2 scfv Ig10 and the scfv molecule with
 CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
 CC for bispecific molecules by cross-pairing the VH and VL domains

CC from the 2 scfvs with different antigen recognition (B7.1/B7.2 and
 CC B7.12/B7.12). The invention relates to novel molecules, including
 CC diabodies, which can cross-link and/or cross-react with the
 CC costimulatory molecules B7.1 and B7.2 expressed on professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. Methods are provided for the
 CC production of such B7-binding molecules, and for their use in the
 CC treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).
 CC
 XX Sequence 273 AA;
 SQ

Query Match 100.0%; Score 48; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
 DB 255 HOYLSSWT 262

RESULT 11
 ID AAW11507 standard; Protein: 301 AA.
 XX AAW11507;
 AC AAW11507;
 XX 24-SEP-1997 (first entry)
 DT
 XX
 DE Single chain, humanised anti-Fc gamma RI antibody H22.
 XX
 XX Humanised antibody; anti-Fc receptor; H22; bifunctional; bispecific;
 KW fusion protein; chimera; carcinoembryonic antigen; CEA.
 XX
 OS Chimeric - Mus musculus.
 OS Chimeric - Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label="signal_peptide"
 FT 20..139
 FT /label="VH"
 FT /note="H22 heavy chain variable region"
 FT 140..154
 FT /label="linker"
 FT 155..266
 FT /label="VL"
 FT /note="H22 light chain variable domain"
 FT 267..281
 FT /label="linker"
 FT 282..292
 FT /label="c-myc_tag"
 FT 296..301
 FT /label="His-6_tail"
 FT /note="6 histidine residues"
 FT
 PN WO9640789-A1.
 PD 19-DEC-1996.
 XX
 PD 07-JUN-1996; 96WO-US09988.
 XX
 PR 07-JUN-1995; 95US-0484172.
 XX
 PA (MEDA-) MEDAREX INC.
 XX
 PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
 DR WPI: 1997-052242/05.
 DR N-PSDB; AAT58129.
 XX

PT Recombinant, multi-specific anti-Fc receptor antibody molecules -
 PT also comprise an anti-target portion, used for the treatment of
 PT cancer, autoimmune disease and pathogenic infection
 XX
 PS Example 8; Fig 39; 115pp; English.
 CC A mammalian expression construct encoding a single chain antibody
 CC having binding specificity for Fc gamma RI, derived from the
 CC humanised anti-Fc gamma RI monoclonal antibody H22 was prepared
 CC (see AAT58129). A bispecific single chain polypeptide was produced
 CC by fusing the H22 scFv to an anti-carcinoembryonic antigen (CEA)
 CC antibody. The H22-anti-CEA fusion protein was shown to bind both
 CC Fc gamma RI and CEA.
 SQ Sequence 301 AA;
 QY
 Db Query Match 100.0%; Score 48; DB 18; Length 301;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HOYLSSWT 8
 |||||
 Db 249 HOYLSSWT 256
 RESULT 12
 AAW73217
 ID AAW73217 standard; Protein: 301 AA.
 XX AAW73217;
 AC
 XX 25-JAN-1999 (first entry)
 DT
 XX
 DE Multispecific single chain antibody H22.
 XX
 KM Multispecific single chain antibody; antibody H22; tumour cell; therapy;
 KM antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
 KM epidermal growth factor receptor; breast cancer; ovarian cancer.
 XX
 OS Homo sapiens.
 XX
 PN US5837243-A.
 PD 17-NOV-1998.
 XX
 XX 07-JUN-1996; 96US-0661052.
 PF
 XX 07-JUN-1996; 96US-0661052.
 PR
 PR 07-JUN-1995; 95US-0484172.
 XX
 PA (MEDA-) MEDAREX INC.
 PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
 XX
 DR WPI: 1999-023374/02.
 DR N-PSDB; AAV08175.
 XX
 PT Specific killing of tumour cells - using a multi-specific molecule
 PT comprising an anti-Fc receptor antibody and a portion which binds to
 PT a target cell
 PS Claim 14; Fig 39; 57pp; English.
 CC This sequence is the multispecific single chain antibody designated
 CC H22. The antibody can be used in the method of the invention for inducing
 CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
 CC which is characterised by overexpression of HER 2/neu or epidermal growth
 CC factor receptor (EGFR), comprises contacting the tumour cell with a
 CC multispecific protein molecule (preferably a single chain antibody)
 CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
 CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
 CC binds to EGFR. The method can be used for treating cancers especially
 CC breast cancer or ovarian cancer. The multispecific antibody can also

CC be administered prophylactically to vaccinate a subject against infection
 CC by a target cell.
 XX
 SQ Sequence 301 AA;
 QY
 Db Query Match 100.0%; Score 48; DB 20; Length 301;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HOYLSSWT 8
 |||||
 Db 249 HOYLSSWT 256
 RESULT 13
 AAB85454
 ID AAB85454 standard; Protein: 301 AA.
 XX AAB85454;
 AC
 XX 25-SEP-2001 (first entry)
 DT
 XX
 DE Single chain humanised anti-Fc gamma RI antibody sequence.
 XX
 KM HER 2/neu; epidermal growth factor receptor; EGFR; multispecific protein;
 KM Fc receptor; FcR; tumour cell; breast; cancer; sarcoma; carcinoma; HIV;
 KM pathogenic; Toxoplasma gondii; candidiasis; systemic lupus; cytostatic;
 KM immune thrombocytopenia purpura; immunosuppressive; antiviral;
 KM antifungal; antiprotzoal.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide
 FT Location/Qualifiers
 FT 1..19
 FT /note= "signal sequence"
 FT 20..139
 FT /note= "H22 Vh"
 FT 140..154
 FT /note= "linker"
 FT 155..266
 FT /note= "H22 Vh"
 FT 267..281
 FT /note= "linker"
 FT 282..292
 FT /note= "c-mys tag"
 FT 296..301
 FT /note= "His-6 tail"
 XX
 XX US6270765-B1.
 PN
 XX
 PD 07-AUG-2001.
 XX
 XX 06-NOV-1998; 98US-0188082.
 PF
 XX 07-JUN-1996; 96US-0661052.
 PR
 PR 07-JUN-1995; 95US-0484172.
 XX
 PA (MEDA-) MEDAREX INC.
 PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
 XX
 DR WPI: 2001-475189/51.
 DR N-PSDB; AAB23382.
 XX
 PT Inducing killing of tumor cells which expresses HER 2/neu or epidermal
 PT growth factor receptor (EGFR) by contacting the cell with multispecific
 PT proteins comprising an anti-Fc receptor, -Her 2/neu or -EGFR antibody,
 PT useful for treating cancer -
 PS Example 8; Fig 39A-B; 57pp; English.
 CC The invention relates to a new method for inducing killing of a tumor
 CC cell which expresses HER 2/neu or epidermal growth factor receptor

CC (EGFR). The method comprises contacting the tumor cell with a
CC multispecific protein comprising a component, preferably an antibody,
CC which binds to an Fc receptor (FcR), Her 2/neu or EGFR. The method is
CC useful for inducing killing of a tumor cell from breast cancer, sarcoma,
CC carcinoma, or ovarian cancer. Specific multispecific proteins can also be
CC administered to a subject to treat or prevent other diseases or
CC conditions, including pathogenic infections (e.g., viral (such as HIV)),
CC protozoan infections (such as Toxoplasma gondii), fungal infections
CC (such as candidiasis), and an autoimmunity (e.g. immune thrombocytopenia
CC purpura and systemic lupus). The present sequence represents the
CC amino acid sequence of a single chain humanised anti-FcgammaRI antibody
CC H22 encoded by construct 225.
XX
SQ Sequence 301 AA:
Query Match 100.0%; Score 48; DB 22; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HOYLSWT 8
DB 249 HOYLSWT 256
RESULT 14
AAB61959
ID AAB61959 standard; Protein: 301 AA.
XX
AC AAB61959;
XX
DT 14-MAY-2001 (first entry)
XX
DE Single chain humanised anti-Fcgamma RI antibody H22.
XX
KM Fcalpha receptor; epidermal growth factor; EGF; HER2 receptor; tumour;
KM immune thrombocytopenia purpura; systemic lupus erythematosus; vaccine;
KM cytostatic; antiviral; protozoicide; antifungal; immunosuppressive;
KM antiinflammatory; dermatological; hemostatic.
XX
OS Mus musculus.
XX
PN WO200109186-A2.
XX
PD 08-FEB-2001.
XX
PE 25-JUL-2000; 2000WO-US20158.
XX
PR 30-JUL-1999; 99US-0364088.
PR 10-MAR-2000; 2000US-0523279.
XX
PA (MEDA-) MEDAREX INC.
XX
PI Deo YM, Goldstein J, Graziano R, Keler T;
XX
DR WPI: 2001-123318/13.
DR N-PSDB; AAC65182.
XX
PT Bispesific molecule comprising specific binding sites for an Fc-alpha
PT receptor and an epidermal growth factor, used to induce effector cell
PT killing of tumor cells -
XX
PS Example 8; Fig 39A-C; 183pp; English.
XX
CC The invention relates to a bispecific molecule (I) comprising specific
CC binding sites for an Fcalpha receptor and an epidermal growth factor
CC (EGF) receptor. It also provides bispecific molecule (II) comprising a
CC human antibody, preferably a single chain antibody, specific for an
CC Fcalpha receptor, linked to EGF; a bispecific molecule (III) comprising
CC specific binding sites for an Fcalpha receptor and a HER2 receptor; (3) a
CC multispecific molecule (IV) comprising specific binding sites for Fcalpha
CC receptor, HER2 receptor and EGF receptor; (4) a multispecific molecule
CC (V) comprising a human antibody specific for an Fcalpha receptor, a human
CC antibody specific for a HER2 receptor, and EGF. (I)-(V) can be used for

CC inducing effector cell killing of tumour cells. The molecules can be used
CC to treat or prevent viral, protozoal, or fungal infections, or autoimmune
CC diseases such as immune thrombocytopenia purpura and systemic lupus
CC erythematosus. The present sequence represents the amino acid sequence of
CC a single chain humanised anti-Fcgamma RI antibody H22.
XX
SQ Sequence 301 AA:
Query Match 100.0%; Score 48; DB 22; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HOYLSWT 8
DB 249 HOYLSWT 256
RESULT 15
AAI06272
ID AAY06272 standard; Protein: 352 AA.
XX
AC AAY06272;
XX
DT 23-AUG-1999 (first entry)
XX
DE Anti-Fc gamma receptor scFv H22-PDGR-R TM fusion.
XX
KM Single chain antibody; scFv; H22; IgG receptor; Fc receptor;
KM Fc gamma receptor; platelet derived growth factor receptor;
KM antibody engineering; cell surface expression; therapy; cancer;
KM tumour; vaccine; human.
XX
XX
OS Chimeric - Homo sapiens.
OS Chimeric - synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..21
FT Peptide /note="Ig K-chain SP"
FT Peptide 22..30
FT Domain /note="HA epitope"
FT Peptide 38..149
FT Peptide /note="H22 VL"
FT Peptide 150..166
FT Domain /note="linker"
FT Peptide 167..286
FT Peptide /note="H22 VH"
FT Peptide 293..302
FT Domain /note="Myc epitope"
FT Domain 303..352
FT Domain /note="PDGFR transmembrane domain"
XX
PN WO928349-A2.
XX
PD 10-JUN-1999.
XX
PE 02-DEC-1998; 98WO-US25556.
XX
PR 02-DEC-1997; 97US-0067232.
XX
PA (MEDA-) MEDAREX INC.
XX
PI Deo YM, Goldstein J, Graziano R, Keler T;
XX
DR WPI: 1999-371099/31.
DR N-PSDB; AAY58935.
XX
PT Cells expressing anti-Fc receptor binding components
XX
PS Example 1; Fig 4A-E; 68pp; English.
XX
CC This sequence represents a fusion protein comprising anti-Fc gamma
CC receptor antibody H22 scFv and the transmembrane domain of platelet
CC derived growth factor receptor (PDGR-R TM). H22 is a humanised

CC monoclonal antibody produced by cell line HA022CL1 (ATCC CRL 11177).
CC The fusion protein is encoded by expression vector pUG717 (see
CC AX58935). The H22 SFV was expressed on the surface of murine tumour
CC cells transformed with this vector. These transformed tumour cells,
CC when in the presence of macrophages, engaged and activated Fc gamma
CC RI, resulting in specific lysis of the tumour cells. Induction of
CC cytokines by H22-TM transformed tumour cells was also demonstrated.
CC This is an example of cells of the invention that have been
CC transformed to express on their surface a component which binds to
CC an Fc receptor of an effector cell. The transformed cell is
CC targeted to an effector cell via the Fc binding component, and can
CC be used as a vehicle to increase an effector cell-mediated immune
CC response, such as cell lysis and phagocytosis, against an antigen
CC associated with the cell. The transformed cells are used to treat
CC cancer and infectious diseases or used as vaccines. The method
CC allows for killing of target cells without targeting any particular
CC antigen on the cell. This is advantageous since many tumour cells
CC and other target cells do not have defined antigens for targeting.
XX

SQ Sequence 352 AA;

Query Match 100.0%; Score 48; DB 20; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQYLSWT 8

DB 132 HQYLSWT 139

Search completed: October 7, 2003, 19:13:54
Job time : 11.069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 4.13793 Seconds
(without alignments)
305.878 Million cell updates/sec

Title: US-09-988-013a-2_COPY_95_102
Perfect score: 48
Sequence: 1 HQTLSWT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCFUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	112	12	US-10-229-335-4
2	48	100.0	112	12	US-10-229-335-28
3	48	100.0	113	10	US-09-741-843-2
4	48	100.0	113	10	US-09-741-843-6
5	48	100.0	113	11	US-09-894-839-2
6	48	100.0	113	11	US-09-894-839-6
7	48	100.0	113	11	US-09-894-839-20
8	48	100.0	113	11	US-09-888-013a-2
9	48	100.0	113	11	US-09-888-013a-6
10	44	91.7	272	12	US-10-053-530-14
11	44	91.7	272	15	US-10-207-655-14
12	39	81.2	112	15	US-10-056-052-10
13	39	81.2	112	15	US-10-056-052-14
14	39	81.2	112	15	US-10-056-052-18
15	37	77.1	154	15	US-10-101-464a-99

16	37	77.1	154	15	US-10-101-464a-768	Sequence 768, App
17	36	75.0	417	12	US-10-342-224-58	Sequence 58, Appl
18	36	75.0	690	15	US-10-101-464a-69	Sequence 69, Appl
19	36	75.0	795	9	US-09-815-242-5499	Sequence 5499, Ap
20	36	75.0	904	9	US-09-815-242-12407	Sequence 12407, A
21	35	72.9	10	9	US-09-813-653-35	Sequence 35, Appl
22	35	72.9	112	15	US-10-056-052-6	Sequence 6, Appl1
23	35	72.9	453	12	US-10-032-585-7887	Sequence 7887, Ap
24	34	70.8	613	15	US-10-013-477-18	Sequence 18, Appl
25	34	70.8	613	15	US-10-013-477-26	Sequence 26, Appl
26	34	70.8	613	15	US-10-207-655-204	Sequence 204, App
27	33	68.8	35	15	US-10-106-698-8240	Sequence 8240, App
28	33	68.8	249	11	US-09-880-748-363	Sequence 363, App
29	33	68.8	578	9	US-09-815-242-12501	Sequence 12501, A
30	33	68.8	886	15	US-10-195-144-61	Sequence 61, Appl
31	32	66.7	70	10	US-09-796-692-1821	Sequence 1821, Ap
32	32	66.7	70	10	US-09-796-692-2076	Sequence 2076, Ap
33	32	66.7	70	15	US-10-040-862-1821	Sequence 1821, Ap
34	32	66.7	70	15	US-10-040-862-2076	Sequence 2076, Ap
35	32	66.7	115	10	US-09-796-692-1443	Sequence 1443, Ap
36	32	66.7	115	10	US-09-796-692-1539	Sequence 1539, Ap
37	32	66.7	115	10	US-09-796-692-1893	Sequence 1893, Ap
38	32	66.7	115	10	US-09-796-692-2021	Sequence 2021, Ap
39	32	66.7	115	15	US-10-040-862-1443	Sequence 1443, Ap
40	32	66.7	115	15	US-10-040-862-1539	Sequence 1539, Ap
41	32	66.7	115	15	US-10-040-862-1893	Sequence 1893, Ap
42	32	66.7	115	15	US-10-040-862-2021	Sequence 2021, Ap
43	32	66.7	235	11	US-09-870-406a-46	Sequence 46, Appl
44	32	66.7	235	15	US-10-159-901-46	Sequence 46, Appl
45	32	66.7	236	11	US-09-870-406a-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-10-229-335-4
Sequence 4, Application US/10229335
Publication No. US20030144483A1

GENERAL INFORMATION:
APPLICANT: MEDAREX, INC.
TITLE OF INVENTION: HUMANIZED ANTIBODIES TO FC RECEPTORS FOR IMMUNOGLOBULIN G ON HUMAN MONONUCLEAR PHAGOCYTES
NUMBER OF SEQUENCES: 28
STREET: P.O. Box 953, 1545 Route 22 East
CITY: Annandale
STATE: New Jersey
COUNTRY: USA
ZIP: 08801

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/229, 335
FILING DATE: 26-Aug-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/435, 516
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35, 430
REFERENCE/DOCKET NUMBER: KAI-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-229-335-4

Query Match
Best Local Similarity 100.0%; Score 48; DB 12; Length 112;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
|||||||
DB 95 HOYLSSWT 102

RESULT 2
US-10-229-335-28
Sequence 28, Application US/10229335
Publication No. US20030144483A1
GENERAL INFORMATION:
APPLICANT: MEDAREX, INC.
TITLE OF INVENTION: HUMANIZED ANTIBODIES TO FC RECEPTORS FOR
IMMUNOBLOBLIN G ON HUMAN MONONUCLEAR PHAGOCYTES
NUMBER OF SEQUENCES: 28
STREET: P.O. Box 953, 1545 Route 22 East
CITY: Annandale
STATE: New Jersey
COUNTRY: USA
ZIP: 08801
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/229,335
FILING DATE: 26-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/435,516
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-10-229-335-28

Query Match
Best Local Similarity 100.0%; Score 48; DB 12; Length 112;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
|||||||
DB 95 HOYLSSWT 102

RESULT 3
US-09-741-843-2
Sequence 2, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans

TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
TITLE OF INVENTION: AND LEUKEMIA CELLS
FILE REFERENCE: 018733/0996
CURRENT APPLICATION NUMBER: US/09/741,843
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-741-843-2

Query Match
Best Local Similarity 100.0%; Score 48; DB 10; Length 113;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
|||||||
DB 95 HOYLSSWT 102

RESULT 4
US-09-741-843-6
Sequence 6, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
TITLE OF INVENTION: AND LEUKEMIA CELLS
FILE REFERENCE: 018733/0996
CURRENT APPLICATION NUMBER: US/09/741,843
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-09-741-843-6

Query Match
Best Local Similarity 100.0%; Score 48; DB 10; Length 113;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
|||||||
DB 95 HOYLSSWT 102

RESULT 5
US-09-894-839-2
Sequence 2, Application US/09894839
Publication No. US20030035800A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
APPLICANT: OU, Zhengxing
TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
FILE REFERENCE: 018733/1049

```

; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-2

Query Match          100.0%; Score 48; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HOYLSWT 8
        |||||||
Db      95 HOYLSWT 102

RESULT 6
US-09-894-839-6
; Sequence 6, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-6

Query Match          100.0%; Score 48; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HOYLSWT 8
        |||||||
Db      95 HOYLSWT 102

RESULT 7
US-09-894-839-20
; Sequence 20, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-20

Query Match          100.0%; Score 48; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HOYLSWT 8
        |||||||
Db      95 HOYLSWT 102

RESULT 8
US-09-988-013a-2
; Sequence 2, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA AND LEUKEMIA CELLS
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-988-013a-2

Query Match          100.0%; Score 48; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HOYLSWT 8
        |||||||
Db      95 HOYLSWT 102

RESULT 9
US-09-988-013a-6
; Sequence 6, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA AND LEUKEMIA CELLS
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
```

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-988-013A-6

Query Match 100.0%; Score 48; DB 11; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
|||
DB 95 HOYLSSWT 102

RESULT 10
US-10-053-530-14
; Sequence 14, Application US/10053530
; Publication No. US20030133939A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey
; APPLICANT: Hayden-Ledbetter, Martha
; TITLE OF INVENTION: Binding Domain-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 390069, 401
; CURRENT APPLICATION NUMBER: US/10/053,530
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: US 09/765,208
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 14
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(272)
; OTHER INFORMATION: MOUSE ANTI-HUMAN CD22 SCFV
US-10-053-530-14

Query Match 91.7%; Score 44; DB 12; Length 272;
Best Local Similarity 87.5%; Pred. No. 5.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
|||
DB 115 HOYLSSWT 122

RESULT 11
US-10-207-655-14
; Sequence 14, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 14
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(272)
; OTHER INFORMATION: MOUSE ANTI-HUMAN CD22 SCFV
US-10-207-655-14

Query Match 91.7%; Score 44; DB 15; Length 272;
Best Local Similarity 87.5%; Pred. No. 5.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
|||
DB 115 HOYLSSWT 122

RESULT 12
US-10-056-052-10
; Sequence 10, Application US/10056052
; Publication No. US2003009656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALI, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLPA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-10

Query Match 81.2%; Score 39; DB 15; Length 112;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
|||
DB 95 HOYLSSWT 102

RESULT 13
US-10-056-052-14
; Sequence 14, Application US/10056052
; Publication No. US2003009656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALI, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLPA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1


```
; SEQ ID NO 14
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-14
```

```
Query Match      81.2%; Score 39; DB 15; Length 112;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HOYLSWT 8
Db      95 HOYLSYT 102
```

```
RESULT 14
US-10-056-052-18
; Sequence 18, Application US/10056052
; Publication No. US20030099636A1
; GENERAL INFORMATION:
; APPLICANT: PATEI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-18
```

```
Query Match      81.2%; Score 39; DB 15; Length 112;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HOYLSWT 8
Db      95 HOYLSYT 102
```

```
RESULT 15
US-10-101-464A-99
; Sequence 99, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
```

```
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-99
```

```
Query Match      77.1%; Score 37; DB 15; Length 154;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HOYLSWT 8
Db      66 HEWLSNWT 73
```

```
Search completed: October 7, 2003, 19:24:25
Job time : 5.13793 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 2.97931 Seconds
(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013a-2_COPY_95_102

Perfect score: 48

Sequence: 1 HOYLSSWT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	81.2	103	2 PH1054	Ig light chain V r
2	39	81.2	111	2 G30502	Ig kappa chain V r
3	38	79.2	686	2 T13490	NADH2 dehydrogenas
4	38	79.2	686	2 T13561	NADH2 dehydrogenas
5	38	79.2	686	2 T13680	NADH2 dehydrogenas
6	38	79.2	688	2 T09836	NADH2 dehydrogenas
7	36	75.0	597	2 A35928	hypothetical 86k p
8	36	75.0	737	2 T28481	hypothetical prote
9	36	75.0	737	2 A72156	C2L protein - vari
10	36	75.0	737	2 T30795	hypothetical prote
11	36	75.0	737	2 F42508	E2L protein - vacc
12	36	75.0	737	2 E36841	conserved hypotet
13	36	75.0	901	2 G89810	hypothetical prote
14	36	75.0	2109	2 T31352	Ig light chain V r
15	35	72.9	101	2 S26337	hypothetical prote
16	35	72.9	174	2 T29783	hypothetical prote
17	35	72.9	381	2 T13666	NADH2 dehydrogenas
18	35	72.9	581	2 A05204	hypothetical prote
19	35	72.9	693	2 T13175	NADH2 dehydrogenas
20	35	72.9	694	2 T13573	NADH2 dehydrogenas
21	35	72.9	694	2 T13572	NADH2 dehydrogenas
22	35	72.9	697	2 T13370	NADH2 dehydrogenas
23	35	72.9	697	2 T13670	NADH2 dehydrogenas
24	35	72.9	698	2 T12625	NADH2 dehydrogenas
25	35	72.9	698	2 T12673	NADH2 dehydrogenas
26	35	72.9	699	2 T12673	NADH2 dehydrogenas
27	35	72.9	701	2 T13587	NADH2 dehydrogenas
28	35	72.9	701	2 T13588	NADH2 dehydrogenas
29	35	72.9	701	2 T13056	NADH2 dehydrogenas

30	35	72.9	702	2 T12624	NADH2 dehydrogenas
31	35	72.9	702	2 T12677	NADH2 dehydrogenas
32	35	72.9	702	2 T13655	NADH2 dehydrogenas
33	35	72.9	702	2 T13409	NADH2 dehydrogenas
34	35	72.9	702	2 T13505	NADH2 dehydrogenas
35	35	72.9	702	2 T13058	NADH2 dehydrogenas
36	35	72.9	703	2 T12696	NADH2 dehydrogenas
37	35	72.9	703	2 T13074	NADH2 dehydrogenas
38	35	72.9	703	2 T13393	NADH2 dehydrogenas
39	35	72.9	703	2 T13696	NADH2 dehydrogenas
40	35	72.9	704	2 T13503	NADH2 dehydrogenas
41	35	72.9	704	2 T13665	NADH2 dehydrogenas
42	35	72.9	705	2 T04400	NADH2 dehydrogenas
43	35	72.9	706	2 T12748	NADH2 dehydrogenas
44	35	72.9	706	2 T13391	NADH2 dehydrogenas
45	35	72.9	734	1 DER2N5	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

PH1054

Ig light chain V region (clone 202.135) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000

C:Accession: PH1054

R:Trillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective

A:Reference number: PH0971; M01D:92381444; PMID:1512540

A:Accession: PH1054

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-103 <RTL>

A:Experimental source: B cell, strain [NZB x NZM]F1

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 81.2% Score 39; DB 2; Length 103;
Best Local Similarity 87.5% Pred. No. 2.6;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
||| ||||| :|
Db 95 HOYLSSWT 102

RESULT 2

G30502

Ig kappa chain V region (A52) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Nov-1988 #sequence_revision 03-Nov-1988 #text_change 21-Jan-2000

C:Accession: G30502

R:Ellat, D.; Webster, D.M.; Rees, A.R.

J. Immunol. 141, 1745-1753, 1988

A:Title: V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F-1

A:Reference number: A30502; M01D:88315787; PMID:2457627

A:Accession: G30502

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-111 <RTL>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 81.2% Score 39; DB 2; Length 111;
Best Local Similarity 75.0% Pred. No. 2.8;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSSWT 8
||: |||||

```
DB      95 HQHFSSWT 102

RESULT 3
T13490
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Mentha x rotundifolia chloroplast
C:Species: chloroplast Mentha x rotundifolia
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Jun-2002
C:Accession: T13490
R:Magstalt, S.J.; Hickerson, L.; Spangler, R.; Reeves, P.A.; Olmstead, R.G.
Plant Syst. Evol. 209, 265-274, 1998
A>Title: Phylogeny of Labiatae s.l. Inferred from cpDNA sequences.
A:Reference number: 217580
A:Accession: T13490
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-686 <MAG>
A:Cross-references: EMBL:U78696; NID:g1695962; PID:g1695963; PIDN:AAB37149.1
C:Genetics:
A:Genome: chloroplast
A>Note: ndhf
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 686;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 HQYLSWT 8
      ||| |||
DB      50 HQYVSWT 57

RESULT 4
T13561
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Plectranthus barbatus chloroplast
C:Species: chloroplast Plectranthus barbatus
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Jun-2002
C:Accession: T13561
R:Magstalt, S.J.; Hickerson, L.; Spangler, R.; Reeves, P.A.; Olmstead, R.G.
Plant Syst. Evol. 209, 265-274, 1998
A>Title: Phylogeny of Labiatae s.l. Inferred from cpDNA sequences.
A:Reference number: 217580
A:Accession: T13561
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-686 <MAG>
A:Cross-references: EMBL:U78698; NID:g1695968; PID:g1695969; PIDN:AAB37152.1
C:Genetics:
A:Genome: chloroplast
A>Note: ndhf
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 686;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 HQYLSWT 8
      ||| |||
DB      50 HQYVSWT 57

RESULT 5
T13680
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Salvia divinorum chloroplast (fr
C:Species: chloroplast Salvia divinorum
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 03-Jun-2002
C:Accession: T13680
R:Magstalt, S.J.; Hickerson, L.; Spangler, R.; Reeves, P.A.; Olmstead, R.G.
Plant Syst. Evol. 209, 265-274, 1998
A>Title: Phylogeny of Labiatae s.l. Inferred from cpDNA sequences.
A:Reference number: 217580

A:Accession: T13680
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-686 <MAG>
A:Cross-references: EMBL:U78703; NID:g1695982; PID:g1695983; PIDN:AAB37159.1
C:Genetics:
A:Genome: chloroplast
A>Note: ndhf
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 686;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 HQYLSWT 8
      ||| |||
DB      50 HQYVSWT 57

RESULT 6
T09836
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - upland cotton chloroplast (fr
C:Species: chloroplast Gossypium hirsutum (upland cotton)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T09836
R:Seelanan, T.; Wendel, J.F.; Schnabel, A.
Syst. Bot. 22, 259-290, 1997
A>Title: Congruence and consensus in the cotton tribe (Malvaceae).
A:Reference number: 216766
A:Accession: T09836
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-688 <SEE>
A:Cross-references: EMBL:U55340; NID:g1654247; PID:g1654248
C:Genetics:
A:Genome: chloroplast
A>Note: ndhf
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 686;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 HQYLSWT 8
      ||| |||
DB      57 HQYVSWT 64

RESULT 7
A35928
hypothetical 86k protein E2L - vaccinia virus
C:Species: vaccinia virus
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 19-Apr-2002
C:Accession: A35928
R:Ahm, B.Y.; Gershon, P.D.; Jones, E.V.; Moss, B.
Mol. Cell. Biol. 10, 5433-5441, 1990
A>Title: Identification of rpo30, a vaccinia virus RNA polymerase gene with structure
A:Reference number: A35928; MUID:90377234; PMID:2398897
A:Accession: A35928
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-597 <AHN>
A:Cross-references: GB:M36339
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match
Best Local Similarity 75.0%; Score 36; DB 2; Length 597;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      2 QYLSWT 8
      :|||||
```

Db 572 RYLSSWT 578

RESULT 8
T28481
hypothetical protein E2L - variola major virus
C:Species: variola major virus
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T28481
R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: Z20488; MUID:9408874; PMID:8264798
A:Accession: T28481
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-737 <MAS>
A:Cross-references: EMBL:U22579; NID:g623595; PIDN:AAA60791.1; PID:g438961
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 75.0%; Score 36; DB 2; Length 737;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSSWT 8
:|||||

Db 572 RYLSSWT 578

RESULT 9
A72156
C2L protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C>Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000
C:Accession: A72156
R:Stchelnikov, S.N.; Tomenin, A.V.; Gutrov, V.V.; Safonov, P.F.; Massung, R.F.; Lopat
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor
A:Reference number: A72150
A:Accession: A72156
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-737 <SHC>
A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54643.1; PID:g5830604
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: C2L
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 75.0%; Score 36; DB 2; Length 737;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSSWT 8
:|||||

Db 572 RYLSSWT 578

RESULT 10
T30795
hypothetical protein 49L - vaccinia virus (strain Ankara)
N:Alternate names: probable 85.9k protein
C:Species: vaccinia virus
A:Variety: strain Ankara
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C:Accession: T30795
R:Antoine, G.; Schellinger, F.; Falkner, F.G.; Dorner, F.
submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A:Reference number: Z20877
A:Accession: T30795
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-737 <ANT>
A:Cross-references: EMBL:U94848; PIDN:AAB96427.1
A:Experimental source: strain Ankara
C:Genetics:
A:Note: MVA049L
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 75.0%; Score 36; DB 2; Length 737;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSSWT 8
:|||||

Db 572 RYLSSWT 578

RESULT 11
F42508
E2L protein - vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus
A:Note: host Homo sapiens (man)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 20-Jun-2000
C:Accession: F42508
R:Johnson, G.P.
submitted to GenBank, June 1990
A:Reference number: A33172
A:Accession: F42508
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-737 <JOH>
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 75.0%; Score 36; DB 2; Length 737;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSSWT 8
:|||||

Db 572 RYLSSWT 578

RESULT 12
E36841
E2L protein - variola virus (strain India-1967)
C:Species: variola virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Mar-2001
C:Accession: E36841
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: E36841
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-737 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA48984.1; PID:g297224
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 75.0%; Score 36; DB 2; Length 737;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSSWT 8
:|||||

Db 572 RYLSSWT 578

RESULT 13
G89810
conserved hypothetical protein SA0412 [imported] - Staphylococcus aureus (strain N315
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89810

R.Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: G89810
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-901 <KUR>
 A:Cross-references: GB:BA000018; PID:g13700344; PIDN:BA041642.1; GSPDB:GN00149
 A:Experimental source: strain NJ15
 C:Genetics:
 A:gene: SA0412

Query Match 75.0%; Score 36; DB 2; Length 901;
 Best Local Similarity 85.7%; Pred. No. 89;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSWT 8
 1:|||||
 Db 177 QYLSWT 183

RESULT 14
 T31352
 hypothetical protein - Pelargonium x hortorum
 C:Species: Pelargonium x hortorum
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T31352
 R:Downie, S.R.; Katz-Downie, D.S.; Wolfe, K.H.; Calle, P.J.; Palmer, J.D.
 Curr. Genet. 25, 367-378, 1994
 A>Title: Structure and evolution of the largest chloroplast gene (ORF280): internal pla
 A:Reference number: 221012; MUID:94363755; PMID:8082181
 A:Accession: T31352
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2109 <DOM>
 A:Cross-references: EMBL:M83200; NID:g468913; PID:g468914; PIDN:AAA73173.1

Query Match 75.0%; Score 36; DB 2; Length 2109;
 Best Local Similarity 75.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HQYLSWT 8
 111111
 Db 28 HQYLSWT 35

RESULT 15
 S26337
 Ig light chain V region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
 C:Accession: S26337; S78449
 R:Stark, S.E.; Caton, A.J.
 J. Exp. Med. 174, 613-624, 1991
 A>Title: Antibodies that are specific for a single amino acid interchange in a protein e
 A:Reference number: S26309; MUID:91341421; PMID:1908510
 A:Accession: S26337
 A:Molecule type: mRNA
 A:Residues: 1-101 <STA>
 A:Cross-references: EMBL:X59193
 R:Caton, A.J.
 submitted to the EMBL Data Library, April 1991
 A:Reference number: S78447
 A:Accession: S78449
 A:Molecule type: mRNA
 A:Residues: 1-60, 'T', 62-91, 'S', 93-101 <CAT>
 A:Cross-references: EMBL:X59193; NID:g52323; PIDN:CAA41903.1; PID:g1334067
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:8-86/Domain: Immunoglobulin homology <IMK>

Query Match 72.9%; Score 35; DB 2; Length 101;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQYLSWT 8
 111111
 Db 87 HQYLSWT 94

Search completed: October 7, 2003, 19:21:48
 Job time : 3.97931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 1.62759 Seconds
(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-2_COPY_95_102
Perfect score: 48
Sequence: 1 HQTLSSWT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwisProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	37	77.1	612	1	PCD8_MOUSE
2	36	75.0	426	1	HEML_SYNEML
3	36	75.0	737	1	VE02_VACCC
4	36	75.0	737	1	VE02_VACCV
5	35	75.0	737	1	VE02_VARY
6	35	72.9	702	1	NU5C_POAPR
7	35	72.9	702	1	NU5C_SORBI
8	35	72.9	705	1	NU5C_HORVU
9	35	72.9	734	1	NU5C_ORISA
10	35	72.9	738	1	NU5C_ORIZA
11	35	72.9	738	1	NU5C_ORIZA
12	35	72.9	738	1	NU5C_ORIZA
13	35	72.9	742	1	NU5C_SPIOL
14	35	72.9	2280	1	YCF2_OENHO
15	34	70.8	238	1	YCF2_TOBAC
16	34	70.8	340	1	SNB2_CHLMU
17	34	70.8	612	1	PCD8_HUMAN
18	34	70.8	613	1	PCD8_HUMAN
19	34	70.8	2131	1	YCF2_SPIOL
20	34	70.8	2216	1	YCF2_SPIOL
21	34	70.8	2294	1	YCF2_SPIOL
22	34	70.8	3066	1	POLG_BCMAN
23	33	68.8	238	1	Y538_CHLTR
24	33	68.8	863	1	AMPN_CAUCR
25	33	68.8	869	1	AMPN_ECOLI
26	33	68.8	869	1	AMPN_ECOLI
27	32	66.7	171	1	AMPN_HAETN
28	32	66.7	172	1	VP19_TBSYB
29	32	66.7	172	1	VP19_TBSYB
30	32	66.7	172	1	VP19_TBSYB
31	32	66.7	172	1	VP19_TBSYB
32	32	66.7	172	1	VP19_TBSYB
33	32	66.7	172	1	VP19_TBSYB

34	32	66.7	220	1	PRT1_PICAN
35	32	66.7	226	1	UNG_VIBCH
36	32	66.7	227	1	UNG_ECOLI
37	32	66.7	228	1	UCRI_SCHPO
38	32	66.7	228	1	UNG_ECO57
39	32	66.7	228	1	UNG_ECOLI
40	32	66.7	228	1	UNG_SALTY
41	32	66.7	236	1	PEBA_SYNPY
42	32	66.7	237	1	DUP1_YEAST
43	32	66.7	241	1	DUC4_HUMAN
44	32	66.7	261	1	TPIS_ENTHI
45	32	66.7	296	1	YDEH_ECOLI
					P31129 escherichia

ALIGNMENTS

RESULT 1
PCD8_MOUSE STANDARD; PRT; 612 AA.
ID PCD8_MOUSE
AC Q920X1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Programmed cell death protein 8, mitochondrial precursor (EC 1.-.-.-)
DE (Apoptosis-inducing factor).
GN PCD8 OR AIF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A., SEQUENCE OF N-TERMINUS, AND SEQUENCE OF 322-336.
RX MEDLINE=99142517; PubMed=9989411;
RA Sustin S.A., Lorenzo H.K., Zamzami N., Marzo I., Snow B.E.,
RA Brothers G.M., Mangion J., Jacotot E., Costantini P., Defflier M.,
RA Larocquette N., Goodlett D.R., Aebersold R., Siderovski D.P.,
RA Penninger J.M., Kroemer G.,
RT "Molecular characterization of mitochondrial apoptosis-inducing factor."
RT Nature 397:441-446(1999).
RL
CC -I- FUNCTION: PROBABLE OXIDOREDUCTASE THAT ACTS AS A CASPASE-
CC INDEPENDENT MITOCHONDRIAL EFFECTOR OF APOPTOTIC CELL DEATH.
CC EXTRAMITOCHONDRIAL AIF INDUCES NUCLEAR CHROMATIN CONDENSATION AND
CC LARGE SCALE DNA FRAGMENTATION (IN VITRO).
CC -I- CORFACTOR: FAD (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Mitochondrial intermembrane space.
CC TRANSLOCATED TO THE NUCLEUS UPON INDUCTION OF APOPTOSIS.
CC -I- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE
CC SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, AF100927, AAD16435.1; -
CC PDB: 1GV4; 31-MAY-02.
CC MGD: MGI:1349419; Pcd8.
CC InterPro: IPR001327; FAD_Pyr_redox.
CC InterPro: IPR001100; Pyr_redox.
CC Pfam: PF00070; Pyr_redox; 1.
CC PRINTS: PR00368; FADPNR.
CC PRINTS: PR00411; PNDPRASE1.
CC OXidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide;
CC Nuclear protein; Apoptosis; 3D-structure.
CC TRANSIT 101
CC CHAIN 102 612
CC DOMAIN 133 438
CC DOMAIN 445 450
CC PROGRAMMED CELL DEATH PROTEIN 8.
CC FAD-DEPENDENT OXIDOREDUCTASE.
CC NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SO SEQUENCE 612 AA; 66765 MW; A17EDP5CF77BB85 CRC64;

Query Match 77.1%; Score 37; DB 1; Length 612;

Best Local Similarity 85.7%; Pred. No. 23;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OYLSWT 8

DB 345 OYLSNMT 351

RESULT 2

HEM1_SYNEL

ID HEM1_SYNEL STANDARD; PRT; 426 AA.

AC Q8D13; 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

GN HEMA OR TLL1738.

OS Synecococcus elongatus (Thermosynechococcus elongatus).

OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.

OX NCBI_TaxID=32046;

RT [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BP-1;

RL MEDLINE-22225144; PubMed-12240834;

RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,

Watanabe A., Iiguchi M., Kawashima K., Kimura T., Kishida Y.,

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,

Shimpo S., Sugimoto M., Takemuchi C., Yamada M., Nakata S.;

RT Complete genome structure of the thermophilic cyanobacterium

Thermosynechococcus elongatus BP-1.";

RL DNA Res. 9:123-130(2002).

CC -1- CATALYTIC ACTIVITY: GlutamyL-trNA(Glu) + NADPH -> glutamate-1-

CC semialdehyde + NADP(+)

CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.

CC Involved in chlorophyll biosynthesis

CC -1- SIMILARITY: Belongs to the glutamyl-trNA reductase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AF005375; BAC09290.1; -

CC HAMAP: MF_00087; -1

CC InterPro: IPR000343; GluTR.

CC InterPro: IPR000594; Thif_domain.

CC Pfam: PF00745; GluTR_dimer; 1.

CC Pfam: PF05201; GluTR_N; 1.

CC Pfam: PF05200; GluTR_NAD_bind; 1.

CC TIGRfam: TIGR01035; hema; 1.

CC PROSITE: PS00747; GluTR; FALSE_NEG.

CC Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;

CC NADP; Complete proteome.

CC ACT_SITE 50 NUCLEOPHILE (BY SIMILARITY).

CC FT ACT_SITE 99 BASE (BY SIMILARITY).

CC SO SEQUENCE 426 AA; 47596 MW; D84CE5A1D2AA777E CRC64;

Query Match 75.0%; Score 36; DB 1; Length 426;

Best Local Similarity 62.5%; Pred. No. 24;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8

DB 70 HOFLSEMS 77

RESULT 3

VE02_VACC ID VE02_VACC STANDARD; PRT; 737 AA.

AC P21080; 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein E2.

GN E2L.

OS Vaccinia virus (strain Copenhagen).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

OC NCBI_TaxID=10249;

RT [1]

RP SEQUENCE FROM N.A.

RL MEDLINE-91021027; PubMed-2219722;

RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,

RA Paolelli E.;

RT "The complete DNA sequence of vaccinia virus.";

RL Virology 179:247-266(1990).

CC [2]

CC COMPLETE GENOME.

RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,

RA Paolelli E.;

RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";

RL Virology 179:517-563(1990).

CC -1- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL: M35027; AAA48039.1; -

CC PIR: F42508; F42508

CC Pfam: PF04497; Fox_E2; 1.

CC SO SEQUENCE 737 AA; 85916 MW; 4DBEB8A1BFC2E2F4 CRC64;

Query Match

Best Local Similarity 85.7%; Score 36; DB 1; Length 737;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OYLSWT 8

DB 572 RYLSWT 578

RESULT 4

VE02_VACC

ID VE02_VACC STANDARD; PRT; 737 AA.

AC P21604; 01-MAY-1991 (Rel. 18, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 01-NOV-1991 (Rel. 20, Last annotation update)

DE Protein E2.

GN E2L.

OS Vaccinia virus (strain WR).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

OC NCBI_TaxID=10254;

RT [1]

RP SEQUENCE FROM N.A.

RL MEDLINE-90377234; PubMed-2398897;

RA Ahn B.-Y., Gershon P.D., Jones E.V., Moss B., Ahn B.-Y.;

RT "Identification of rpo30, a vaccinia virus RNA polymerase gene with

RT structural similarity to a eucaryotic transcription elongation

RT factor.";


```

RL MOL. Cell. Biol. 10:5433-5441(1990).
CC -1- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M36339; AAB59822.1; -.
DR PIR: A35928; A35928.
DR Pfam: PF04497; Pox_E2; 1.
SQ SEQUENCE 737 AA; 85957 MW; D5BAF09E2F944A9E CRC64;

Query Match 75.0%; Score 36; DB 1; Length 737;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYLSWT 8
Db 572 RYLSWT 578

RESULT 5
VE02_VARY STANDARD; PRT; 737 AA.
AC P33862;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein E2.
GN E2L.
OS Variola virus.
OC Viruses: dsDNA viruses, no RNA stage: Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC NCBI_TaxID=10255;
OX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-India-1967 / Isolate Ind3;
RX MEDLINE=94152154; PubMed=8109158;
RA Sandakchlev L.S.;
RA Sandakchlev S.N., Blinov V.M., Resenchuk S.M., Totmenin A.V.,
RT "Analysis of the nucleotide sequence of a 43 kbp segment of the
RT genome of variola virus India-1967 strain.";
RL Virus Res. 30:239-258(1993).
RN [2]
RP COMPLETE GENOME.
RC STRAIN-India-1967 / Isolate Ind3;
RX MEDLINE=93202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakchlev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
RN -1- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69198; CAA48984.1; -.
DR PIR: E36841; E36841.
DR Pfam: PF04497; Pox_E2; 1.
SQ SEQUENCE 737 AA; 85957 MW; F8CA3A5A6EFA17B0 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 737;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 QYLSWT 8
Db 572 RYLSWT 578

RESULT 6
NU5C_POAPR STANDARD; PRT; 702 AA.
ID NU5C_POAPR
AC Q32880;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase
DE chain 5) (Fragment).
GN NDH-F.
OS Poa pratensis (Kentucky bluegrass).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Poaceae; Poa.
OX NCBI_TaxID=4545;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Clark L.G., Zhang W., Wendel J.F.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC plastoquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U21980; AAA64698.1; -.
DR PIR: T13655; T13655.
DR InterPro: IPR003916; NADHdb_oxrds.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS: PR01434; NADHDGNASES.
DR Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
FT NON_TER 1
FT NON_TER 702
SQ SEQUENCE 702 AA; 78718 MW; 7D7E2C6AF961185F CRC64;

Query Match 72.9%; Score 35; DB 1; Length 702;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
Db 61 YOYLSWT 68

RESULT 7
NU5C_SORBI STANDARD; PRT; 702 AA.
ID NU5C_SORBI
AC Q33066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase
DE chain 5) (Fragment).

```

```
GN NDHF.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Clark L.G., Zhang W., Wendel J.F.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastocquinone = NAD(P)(+) +
CC plastocquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U21981; AAA64699.1; -
CC InterPro: IPR003916; NADhub_oxred5.
CC InterPro: IPR001750; Oxidored_q1.
CC InterPro: IPR002128; Oxidored_q1_C.
CC InterPro: IPR001516; Oxidored_q1_N.
CC Pfam: PF00361; Oxidored_q1; 1.
CC Pfam: PF01010; Oxidored_q1_C; 1.
CC Pfam: PF00662; Oxidored_q1_N; 1.
CC PRINTS: PR01434; NADHDGNASE5.
CC Oxidoreductase; NAD; NADP; Quinone; Plastocquinone; Chloroplast.
FT NON_TER 1 702
FT SEQUENCE 702 AA; 78504 MW; 186BEAF2F956F851 CRC64;
SQ
Query Match 72.9%; Score 35; DB 1; Length 702;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 HOYLSSWT 8
DB 61 YQYLMSWT 68
RESULT 8
NU5C_HORVU STANDARD; PRT; 705 AA.
ID NU5C_HORVU 032440;
AC 032440;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastocquinone oxidoreductase
DE chain 5) (Fragment).
GN NDHF.
OS Hordeum vulgare (Barley).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Clark L.G., Zhang W., Wendel J.F.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastocquinone = NAD(P)(+) +
CC plastocquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U22003; AAA64207.1; -
CC PIR: T04400; T04400.
CC InterPro: IPR003916; NADhub_oxred5.
CC InterPro: IPR001750; Oxidored_q1.
CC InterPro: IPR002128; Oxidored_q1_C.
CC InterPro: IPR001516; Oxidored_q1_N.
CC Pfam: PF00361; Oxidored_q1; 1.
CC Pfam: PF01010; Oxidored_q1_C; 1.
CC Pfam: PF00662; Oxidored_q1_N; 1.
CC PRINTS: PR01434; NADHDGNASE5.
CC Oxidoreductase; NAD; NADP; Quinone; Plastocquinone; Chloroplast.
FT NON_TER 1 705
FT SEQUENCE 705 AA; 78841 MW; ACC63E07C070ADE6 CRC64;
SQ
Query Match 72.9%; Score 35; DB 1; Length 705;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 HOYLSSWT 8
DB 61 YQYLMSWT 68
RESULT 9
NU5C_ORYSA STANDARD; PRT; 734 AA.
ID NU5C_ORYSA P12129;
AC P12129;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastocquinone oxidoreductase
DE chain 5).
GN NDHF.
OS Oryza sativa (Rice).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC MEDLINE=89364698; PubMed=2770692;
RA Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M.,
RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Suglura M.;
RT "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RT intermolecular recombination between distinct tRNA genes accounts for
RT a major plastid DNA inversion during the evolution of the cereals.";
RT Mol. Gen. Genet. 217:185-194(1989).
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastocquinone = NAD(P)(+) +
CC plastocquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15901; CAA33950.1; -
CC PIR: J00286; DERZNS.
CC Gramene; P12129;
CC InterPro: IPR003916; NADhub_oxred5.
```

DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS: PR01434; NADHGNASE5.
KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 734 AA; 82597 MW; BFEFE20A01632B2 CRC64;

Query Match 72.9%; Score 35; DB 1; Length 734;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 69 YOYLSWT 76

RESULT 10
NUSC_MAIZE STANDARD; PRT; 738 AA.
ID NUSC_MAIZE
AC P46620;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase chain 5).
GN NDHF OR NDH5.
OS Zea mays (Maize).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RX MEDLINE=95395841; PubMed=7666415;
RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing.";
RT J. Mol. Biol. 251:614-628(1995).
RL [2]
RN SEQUENCE OF 9-709 FROM N.A.
RP TISSUE=Leaf;
RA Clark L.G., Zhang W., Wendel J.F.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) + plastoquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85653; CAA60346.1; -;
DR EMBL; U21985; AAA64703.1; -;
DR PIR; S58612; S58612.
DR MaltzEB: 107782; -;
DR InterPro: IPR003916; NADhub_oxred5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHGNASE5.
KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.

FT CONFLICT 268 268 MISSING (IN REF. 2).
FT CONFLICT 289 289 W -> L (IN REF. 2).
FT CONFLICT 570 570 R -> G (IN REF. 2).
FT CONFLICT 680 680 K -> R (IN REF. 2).
SQ SEQUENCE 738 AA; 82976 MW; F4E3EBDDDC91FA CRC64;

Query Match 72.9%; Score 35; DB 1; Length 738;
Best Local Similarity 75.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 69 YOYLSWT 76

RESULT 11
NUSC_WHEAT STANDARD; PRT; 739 AA.
ID NUSC_WHEAT
AC Q95H46;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase chain 5).
GN NDHF.
OS Triticum aestivum (Wheat).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Chinese Spring;
RA Ogihara Y., Isono K., Kojima T., Endo A., Hanaoka M., Shina T.,
RA Terachi T., Utsugi S., Murata M., Mori N., Takumi S., Ikeo K.,
RA Gojobori T., Mural R., Mural K., Matsunaka Y., Ohnishi Y., Tajiri H.,
RA Tsunewaki K.;
RT "Chinese spring wheat (Triticum aestivum L.) chloroplast genome: complete sequence and contig clones.";
RT Plant Mol. Biol. Rep. 18:243-253(2000).
RL [1]
RN CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) + plastoquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB042240; BAB47082.1; -;
DR InterPro: IPR003916; NADhub_oxred5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHGNASE5.
KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 739 AA; 82860 MW; D22C99E699A3354F CRC64;

Query Match 72.9%; Score 35; DB 1; Length 739;
Best Local Similarity 75.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 69 YOYLSWT 76

```

RESULT 12
NCBI_SPTOL
ID NCBI_SPTOL STANDARD: PRT: 742 AA.
AC 09M3J4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 5, chloroplast (RC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase
  chain 5).
GN NDHF.
OS Spinacia oleracea (Spinach).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Geant d'herve, and cv. Monatol;
RC MEDLINE=21187424; PubMed=11292076;
RA Schmitz-Linneberger C., Maier R.M., Alcaraz J.-P., Cottet A.,
RA Hermann R.G., Maché R.;
RT "The plastid chromosome of spinach (Spinacia oleracea): complete
  nucleotide sequence and gene organization."
RT Plant Mol. Biol. 45:307-315(2001).
RL -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone -> NAD(P)(+) +
  plastoquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ400848; CAB88780.1; -
DR InterPro: IPR003916; NADHdb_oxred5.
DR InterPro: IPR001750; Oxidored_g1.
DR InterPro: IPR002128; Oxidored_g1_C.
DR InterPro: IPR001516; Oxidored_g1_N.
DR Pfam: PF00361; oxidored_g1; 1.
DR Pfam: PF01010; oxidored_g1_C; 1.
DR Pfam: PF00662; oxidored_g1_N; 1.
DR PRINTS: PR01434; NADHGNAS5.
DR OXIDOREDUCTASE: NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 742 AA; 84529 MW; AB84C8A15E185E7F CRC64;

Query Match
Best Local Similarity 72.9%; Score 35; DB 1; Length 742;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
DB 69 YOYLSWT 76

RESULT 13
YCF2_OENHO
ID YCF2_OENHO STANDARD: PRT: 2280 AA.
AC 09MEF2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 266.3 kDa protein ycf2.
GN YCF2-A AND YCF2-B.
OS Oenothera hookeri (Hooker's evening primrose).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.

```

```

OX NCBI_TaxID=85636;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Johansen;
RC MEDLINE=20309318; PubMed=10852478;
RA Hupfer H., Swiatek M., Hornung S., Hermann R.G., Maier R.M.,
RA Chlu W.-L., Sears B.;
RT "Complete nucleotide sequence of the Oenothera elata plastid
  chromosome, representing plastome I of the five distinguishable
  RT Eucenothera plastomes."
RL Mol. Gen. Genet. 263:581-585(2000).
CC -1- SIMILARITY: BELONGS TO THE YCF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ271079; CAB67203.1; -
DR EMBL: AJ271079; CAB67242.1; -
DR InterPro: IPR003593; AAA_Arpase.
DR SMART: SM00382; AAA; 1.
KM Chloroplast; Hypothetical protein.
SQ SEQUENCE 2280 AA; 266257 MW; 0DC3FD2EF9358657 CRC64;

Query Match
Best Local Similarity 72.9%; Score 35; DB 1; Length 2280;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
DB 24 HHFLDSWT 31

RESULT 14
YCF2_TOBAC
ID YCF2_TOBAC STANDARD: PRT: 2280 AA.
AC P09976; P09977;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 267 kDa protein ycf2 (ORF 2280).
GN YCF2.
OS Nicotiana tabacum (Common tobacco).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bright yellow 4;
RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,
RA Matsubayashi T., Zaita N., Chunwongse J., Ohkita J.,
RA Yamaguchi-Shinozaki K., Ohno C., Torazawa K., Meng B.-Y., Sugita M.,
RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,
RA Tohoh N., Shimada H., Sugita M.;
RT "The complete nucleotide sequence of the tobacco chloroplast genome:
  RT its gene organization and expression."
RL EMBO J. 5:2043-2049(1986).
RN 12
RP REVISIONS.
RA Sugita M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NOT YET KNOWN.
CC -1- SIMILARITY: BELONGS TO THE YCF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z00044: CAA77427.1; -
DR EMBL: Z00044: CAA77438.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase-centr.
DR Pfam: PF00004; AAA; 1.
DR SMART: SM00382; AAA; 1.
KM Chloroplast; Hypothetical protein.
SQ SEQUENCE 2280 AA; 266812 MW; E246D5F3D902C06D CRC64;

Query Match 72.9%; Score 35; DB 1; Length 2280;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HOYLSSWT 8
| : | | |
DB 24 HHFLDSWT 31

RESULT 15
Y825_CHLMU STANDARD; PRT; 238 AA.
ID Y825_CHLMU
AC O9PJRK4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TC0825.
GN TC0825.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MOPn / N199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0658/CT538/TC0825
CC FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE002349; AAF39626.1; -
DR PIR: B81660; B81660.
DR TIGR: TC0825; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 238 AA; 27303 MW; 5FF62FD7D893D049 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 238;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSSW 7
| : | | |
DB 227 HEYISQW 233

Search completed: October 7, 2003, 19:14:59
Job time : 2.62759 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 7.97241 Seconds
(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-2_COPY_95_102
Perfect score: 48
Sequence: 1 HOYLSSWT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	41	85.4	658	8 O9GDW8	O9gdw8 androcymbin
2	39	81.2	686	8 P92299	P92299 glechoma he
3	39	81.2	705	8 O8M9R0	O8m9r0 eucommia ul
4	39	81.2	730	8 O8HTN9	O8htn9 akebia quhn
5	39	81.2	3202	12 O8QY01	O8qy01 bean common
6	38	79.2	610	8 O9BBS5	O9bbs5 maschaloccp
7	38	79.2	657	8 O9GDV8	O9gdv8 tricyrtis l
8	38	79.2	660	8 O9GDV9	O9gdv9 tricyrtis a
9	38	79.2	686	8 P92344	P92344 salvia divi
10	38	79.2	686	8 O03646	O03646 mentha x ro
11	38	79.2	686	8 P92329	P92329 plectranthu
12	38	79.2	688	8 O95645	O95645 gossypium h
13	38	79.2	701	8 O8HTM4	O8htm4 achlys trip
14	38	79.2	702	10 O9SC10	O9sc10 tetracera a
15	38	79.2	704	8 O8HTM2	O8htm2 leonlice ev
16	38	79.2	704	8 O8HTM1	O8htm1 gymnosperm

17	38	79.2	726	8 O8HTN2	O8htn2 ranzania ja
18	38	79.2	726	8 O8HTM9	O8htm9 sinopodophy
19	38	79.2	726	8 O8HTM8	O8htm8 diphyllaia
20	38	79.2	726	8 O8HTM7	O8htm7 podophyllum
21	38	79.2	726	8 O8HTM6	O8htm6 dysosma ple
22	38	79.2	728	8 O8HTN4	O8htm4 nandina dom
23	38	79.2	728	8 O8HTN3	O8htm3 caulophyllu
24	38	79.2	728	8 O8HTM5	O8htm5 bongardia c
25	38	79.2	728	8 O8HTL9	O8htl9 eplmedium k
26	38	79.2	729	8 O8HTN1	O8htn1 jeffersonia
27	38	79.2	729	8 O8HTN0	O8htn0 jeffersonia
28	37	77.1	429	8 O9GFR3	O9gfr3 liriiodendro
29	37	77.1	731	8 O95D45	O95d45 liriiodendro
30	37	77.1	732	8 O95D44	O95d44 liriiodendro
31	37	77.1	732	8 O9TL43	O9tl43 liriiodendro
32	37	77.1	743	8 O9GE24	O9gez4 gilia rigid
33	37	77.1	853	16 O92MT9	O92mt9 rhizobium m
34	36	75.0	145	10 O94DT5	O94dt5 oryza sativ
35	36	75.0	409	16 O8KC60	O8kc60 chlorobium
36	36	75.0	426	16 O8DI53	O8di53 synchococc
37	36	75.0	737	12 O8OMZ8	O8omz8 cowpox viru
38	36	75.0	737	12 O8O371	O8o371 variola maj
39	36	75.0	737	12 O8PX52	O8px52 variola vir
40	36	75.0	737	12 O8ULG3	O8ulg3 ectromelia
41	36	75.0	737	12 O9QNK3	O9qnk3 variola min
42	36	75.0	737	12 O57185	O57185 vaccinia vi
43	36	75.0	737	12 O8Y2W3	O8y2w3 camelpox vi
44	36	75.0	737	12 O9JFD8	O9jfd8 vaccinia vi
45	36	75.0	843	10 O8H821	O8h821 oryza sativ

ALIGNMENTS

RESULT 1

Q9GDW8 PRELIMINARY; PRT; 658 AA.

AC O9GDW8;
ID O9GDW8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit F (Fragment).
GN NDHF.
OS Androcymbium ciliolatum.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Colchicaceae;
OC Androcymbium.
OX NCBI_TaxID=59023;
RN [1]
RP SEQUENCE FROM N.A.
RA Patterson T.B., Glynnish T.J.;
RT "Phylogeny, concerted convergence, and phylogenetic niche conservatism
in the core Liliales: Insights from rbcL and ndhF sequence data.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276012; AAC25128.1; -;
DR InterPro: IPR003916; NADHdh_oxred5.
DR InterPro: IPR001750; Oxidored_g1.
DR InterPro: IPR002128; Oxidored_g1_C.
DR InterPro: IPR001516; Oxidored_g1_N.
DR Pfam: PF00361; Oxidored_g1; 1.
DR Pfam: PF01010; Oxidored_g1_C; 1.
DR Pfam: PF00662; Oxidored_g1_N; 1.
DR PRINTS; PR01434; NADHdHGNASE5.
KW NAD; Oxidoreductase; Plastiquinone; Chloroplast.
FT NON_TER 1
FT NON_TER 658
SQ SEQUENCE 658 AA; 74621 MW; CEEA0AB25AC81880 CRC64;

Query Match 85.4%; Score 41; DB 8; Length 658;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
 |||||
 DB 40 HOYLSWT 47

RESULT 2

ID P92299 PRELIMINARY; PRT; 686 AA.
 AC P92299;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE NADH dehydrogenase (Fragment).
 GN NDHF.
 OS Glechoma hederacea (Ground-ivy).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Lamiaceae; Nepetoideae; Menthaceae;
 OC Glechoma.
 OC NCBI_TaxID=28509;
 OX [1]
 RN
 RP SEQUENCE FROM N.A.
 RA Wagstaff S.J., Hickerson L., Spangler R., Reeves P.A., Olmstead R.G.;
 RT "Phylogeny of Labiatae s.l. inferred from cpDNA sequences.";
 RL Plant Syst. Evol. 0:0-0(1997).
 DR EMBL: U78691; AAB37144.1; -;
 DR InterPro: IPR001064; Crystal11n.
 DR InterPro: IPR003916; NADhub_oxred5.
 DR InterPro: IPR001750; Oxidored_q1_C.
 DR InterPro: IPR002128; Oxidored_q1_C.
 DR InterPro: IPR001516; Oxidored_q1_N.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR Pfam: PF01010; oxidored_q1_C; 1.
 DR Pfam: PF00662; oxidored_q1_N; 1.
 DR PRINTS: PR01434; NADHDGNASE5.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 DR NAD: Oxidoreductase; Plastocquinone; Chloroplast.
 KM NAD: Oxidoreductase; Plastocquinone; Chloroplast.
 FT NON_TER 686
 SQ SEQUENCE 686 AA; 77786 MW; BA6381A5DC0BAC32 CRC64;

Query Match 81.2%; Score 39; DB 8; Length 686;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
 |||||
 DB 50 HOYLSWT 57

RESULT 3

ID Q8M9R0 PRELIMINARY; PRT; 705 AA.
 AC Q8M9R0;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE NADH dehydrogenase subunit F (Fragment).
 GN NDHF.
 OS Eucommia ulmoides.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Garryales; Eucommiaceae; Eucommia.
 OC NCBI_TaxID=4392;
 OX [1]
 RN
 RP SEQUENCE FROM N.A.
 RA Bremer B., Bremer K., Heldart N., Erikson P., Olmstead R.G.,
 RA Anderberg A.A., Kallersjo M., Barkhofdarian E.;
 RT "Phylogenetic analysis of asterids based on 3 non-coding
 RT chloroplast DNA markers and the utility of non-coding DNA at higher
 RT taxonomic levels.";
 RL Mol. Phylogenet. Evol. 24:273-300(2002).

DR EMBL: AJ429113; CAD23091.1; -;
 DR InterPro: IPR003916; NADhub_oxred5.
 DR InterPro: IPR001750; Oxidored_q1_C.
 DR InterPro: IPR002128; Oxidored_q1_C.
 DR InterPro: IPR001516; Oxidored_q1_N.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR Pfam: PF01010; oxidored_q1_C; 1.
 DR Pfam: PF00662; oxidored_q1_N; 1.
 DR PRINTS: PR01434; NADHDGNASE5.
 DR NAD: Oxidoreductase; Plastocquinone; Chloroplast.
 KM NAD: Oxidoreductase; Plastocquinone; Chloroplast.
 FT NON_TER 705
 SQ SEQUENCE 705 AA; 79613 MW; 7703F997E04543FF CRC64;

Query Match 81.2%; Score 39; DB 8; Length 705;
 Best Local Similarity 75.0%; Pred. No. 56;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
 |||||
 DB 61 YOYVSSWT 68

RESULT 4

ID Q8HTN9 PRELIMINARY; PRT; 730 AA.
 AC Q8HTN9;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE NADH dehydrogenase subunit F (Fragment).
 GN NDHF.
 OS Akebia quinata (Five leaf akebia).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 OC Lardizabalaceae; Akebia.
 OC NCBI_TaxID=13331;
 OX [1]
 RN
 RP SEQUENCE FROM N.A.
 RA Kim Y.-D., Kim S.-H., Jansen R.K.;
 RT "Phylogeny of the Berberidaceae based on the sequences of the
 RT chloroplast gene ndhF.";
 RL Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY145143; AAN64386.1; -;
 KM Chloroplast.
 FT NON_TER 730
 SQ SEQUENCE 730 AA; 81618 MW; CAPE1770BF1B3A75 CRC64;

Query Match 81.2%; Score 39; DB 8; Length 730;
 Best Local Similarity 75.0%; Pred. No. 58;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
 |||||
 DB 69 YOYVSSWT 76

RESULT 5

ID Q8QY01 PRELIMINARY; PRT; 3202 AA.
 AC Q8QY01;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE Polyprotein.
 OS Bean common mosaic virus (strain NL-3 / Michigan) (BCMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OC NCBI_TaxID=12196;
 OX [1]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-blackeye cowpea mosaic;

RA Adams M.J.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-blackeye cowpea mosaic;
 RA Zheng H.Y., Chen J., Chen J.P., Adams M.J., Hou M.S.;
 RT "bean common mosaic virus isolates causing different symptoms in asparagus bean in China differ greatly in the 5'-parts of their genomes";
 RL Arch. Virol. 147:0-0(2002).
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
 DR EMBL: AJ12437; CAC86160.1; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001456; Peptidase_C6.
 DR InterPro: IPR001592; poly-coat.
 DR InterPro: IPR002540; poly_pl.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_PSVLR.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF00863; Peptidase_C4; 1.
 DR Pfam: PF00851; Peptidase_C6; 1.
 DR Pfam: PF00767; poly-coat; 1.
 DR Pfam: PF01577; poly_pl; 1.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICC; 1.
 DR PROSITE: PSS0507; RDRP_POSITIVE; 1.
 DR PROSITE: PSS0521; RDRP_VIRAL; 1.
 KW ATP-binding; Coat protein; Helicase; Hydrolase.
 FT CHAIN 1 423
 FT CHAIN 424 880
 FT CHAIN 881 1227
 FT CHAIN 1228 1279
 FT CHAIN 1280 1913
 FT CHAIN 1914 1966
 FT CHAIN 1967 2156
 FT CHAIN 2157 2399
 FT CHAIN 2400 2915
 FT CHAIN 2916 3202
 SQ SEQUENCE 3202 AA; 363915 MW; 2BA72C32BCD65548 CRC64;

Query Match 81.2%; Score 39; DB 12; Length 3202;
 Best Local Similarity 75.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
 DB 1743 HOYLSWT 1750

RESULT 6
 O9BB58 PRELIMINARY; PRT; 610 AA.
 AC O9BB58;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE MDH dehydrogenase F (Fragment).
 GN MDH.
 OS Maschalocephalus dinklagei.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Commelinidae; Insectae; sedis;
 CC Rapateaceae; Maschalocephalus.
 CX NCBI_TaxID=142432;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Glyvish T.J., Evans T.M., Zjhra M.L., Patterson T.B., Berry P.E.,
 RA Sysma K.J.;
 RT "Molecular evolution, adaptive radiation, and geographic

RT diversification in the amphitlantic family Rapateaceae: evidence from
 RT ndyf sequences and morphology";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF207628; AAK21836.1; -.
 DR InterPro: IPR003916; NADHub_oxred5.
 DR InterPro: IPR001750; Oxidored_q1.
 DR InterPro: IPR002128; Oxidored_q1_C.
 DR InterPro: IPR001516; Oxidored_q1_N.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR Pfam: PF01010; oxidored_q1_C; 1.
 DR Pfam: PF00662; oxidored_q1_N; 1.
 DR PRINTS: PRO1434; NADHGNASE5.
 KW NAD; Oxidoreductase; Plastoquinone; Chloroplast.
 FT NON_TER 1 610
 FT NON_TER 610 610
 SQ SEQUENCE 610 AA; 68669 MW; 3AF1F9E61C086AD1 CRC64;

Query Match 79.2%; Score 38; DB 8; Length 610;
 Best Local Similarity 75.0%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
 DB 7 YOYLSWT 14

RESULT 7
 O9GDV8 PRELIMINARY; PRT; 657 AA.
 AC O9GDV8;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE MDH dehydrogenase subunit F (Fragment).
 GN MDH.
 OS Tricyrtis latifolia.
 OC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Liliaceae; Colchicaceae;
 CC Tricyrtis.
 CX NCBI_TaxID=85291;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Patterson T.B., Glyvish T.J.;
 RT "Phylogeny, concerted convergence, and phylogenetic niche conservatism in the core Liliaceae: Insights from rbcL and ndyf sequence data";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF276022; AAC25138.1; -.
 DR InterPro: IPR003916; NADHub_oxred5.
 DR InterPro: IPR001750; Oxidored_q1.
 DR InterPro: IPR002128; Oxidored_q1_C.
 DR InterPro: IPR001516; Oxidored_q1_N.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR Pfam: PF01010; oxidored_q1_C; 1.
 DR Pfam: PF00662; oxidored_q1_N; 1.
 DR PRINTS: PRO1434; NADHGNASE5.
 KW NAD; Oxidoreductase; Plastoquinone; Chloroplast.
 FT NON_TER 1 657
 FT NON_TER 657 657
 SQ SEQUENCE 657 AA; 73991 MW; 13DC69801A1C360A CRC64;

Query Match 79.2%; Score 38; DB 8; Length 657;
 Best Local Similarity 75.0%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
 DB 40 YOYLSWT 47

RESULT 8
 O9GDV9 PRELIMINARY; PRT; 660 AA.
 ID O9GDV9

```

AC 09GDV9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit F (Fragment).
GN NDHF.
OS Tricyrtis affinis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Colchicaceae;
OC Tricyrtis.
OX NCBI_TaxID-34198;
RN [1]
RP SEQUENCE FROM N.A.
RA Paterson T.B., Givnish T.J.;
RT *Phylogeny, concerted convergence, and phylogenetic niche conservatism
RT in the core Liliales: Insights from rbcL and ndhF sequence data.";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF276021.1;
DR InterPro: IPR003916; NADhub.oxted5.
DR InterPro: IPR001750; Oxidored_g1_C.
DR InterPro: IPR002128; Oxidored_g1_C.
DR InterPro: IPR001516; Oxidored_g1_N.
DR Pfam: PF00361; Oxidored_g1_C.1.
DR Pfam: PF01010; Oxidored_g1_C.1.
DR Pfam: PF00662; Oxidored_g1_N.1.
DR PRINTS: PR01434; NADHDHGNAS5.
DR NAD; Oxidoreductase; Plastocquinone; Chloroplast.
FT NON_TER 1
FT SEQUENCE 660 AA; 74316 MW; 7E549CCAC6DD5CCB CRC64;
SQ
Query Match 79.2%; Score 38; DB 8; Length 660;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
DB 40 HOYLSWT 47

RESULT 9
P92344 PRELIMINARY; PRT; 686 AA.
AC P92344;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Salvia divinorum (Maria pastora).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Lamiaceae; Nepetoideae; Mentheae;
OC Salvia.
OX NCBI_TaxID-28513;
RN [1]
RP SEQUENCE FROM N.A.
RA Magataff S.J., Hickerson L., Spangler R., Reeves P.A., Olmstead R.G.;
RT "Phylogeny of Labiatae s.l. inferred from cpDNA sequences.";
RL Plant Syst. Evol. 0:0-0(1997).
DR EMBL: U78703; AAB37159.1;
DR InterPro: IPR001064; Crystal1in.
DR InterPro: IPR003916; NADhub.oxted5.
DR InterPro: IPR001750; Oxidored_g1_C.
DR InterPro: IPR002128; Oxidored_g1_C.
DR InterPro: IPR001516; Oxidored_g1_N.
DR Pfam: PF00361; Oxidored_g1_C.1.
DR Pfam: PF01010; Oxidored_g1_C.1.
DR Pfam: PF00662; Oxidored_g1_N.1.
DR PRINTS: PR01434; NADHDHGNAS5.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR NAD; Oxidoreductase; Plastocquinone; Chloroplast.
FT NON_TER 686
FT SEQUENCE 686 AA; 77675 MW; 72C799DB30C92020 CRC64;
SQ
Query Match 79.2%; Score 38; DB 8; Length 686;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
DB 50 HOYLSWT 57

RESULT 11
P92329 PRELIMINARY; PRT; 686 AA.
AC P92329;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Plectranthus barbatus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Lamiaceae; Nepetoideae; Ocimeae;
OC Plectranthus.

```

```

KW NAD; Oxidoreductase; Plastocquinone; Chloroplast.
FT NON_TER 686
FT SEQUENCE 686 AA; 77863 MW; FBFD32F3A9336AB CRC64;
SQ
Query Match 79.2%; Score 38; DB 8; Length 686;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
DB 50 HOYLSWT 57

RESULT 10
O03646 PRELIMINARY; PRT; 686 AA.
AC O03646;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Mentha x rotundifolia.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Lamiaceae; Nepetoideae; Mentheae;
OC Mentha.
OX NCBI_TaxID-54461;
RN [1]
RP SEQUENCE FROM N.A.
RA Magataff S.J., Hickerson L., Spangler R., Reeves P.A., Olmstead R.G.;
RT "Phylogeny of Labiatae s.l. inferred from cpDNA sequences.";
RL Plant Syst. Evol. 0:0-0(1997).
DR EMBL: U78696; AAB37149.1;
DR InterPro: IPR001064; Crystal1in.
DR InterPro: IPR003916; NADhub.oxted5.
DR InterPro: IPR001750; Oxidored_g1_C.
DR InterPro: IPR002128; Oxidored_g1_C.
DR InterPro: IPR001516; Oxidored_g1_N.
DR Pfam: PF00361; Oxidored_g1_C.1.
DR Pfam: PF01010; Oxidored_g1_C.1.
DR Pfam: PF00662; Oxidored_g1_N.1.
DR PRINTS: PR01434; NADHDHGNAS5.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR NAD; Oxidoreductase; Plastocquinone; Chloroplast.
FT NON_TER 686
FT SEQUENCE 686 AA; 77675 MW; 72C799DB30C92020 CRC64;
SQ
Query Match 79.2%; Score 38; DB 8; Length 686;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HOYLSWT 8
DB 50 HOYLSWT 57

RESULT 11
P92329 PRELIMINARY; PRT; 686 AA.
AC P92329;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Plectranthus barbatus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Lamiaceae; Nepetoideae; Ocimeae;
OC Plectranthus.

```

OX NCBI_TaxID=41228;
[1]
RP SEQUENCE FROM N.A.
RA Magstiff S.J., Hickerson L., Spangler R., Reeves P.A., Olmstead R.G.;
RT "Phylogeny of Labiales s.l. inferred from cpDNA sequences."
RL Plant Syst. Evol. 0:0-0(1997).
DR EMBL: U78698; AAB37152.1; -
DR InterPro: IPR003916; NADHdb_oxred5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS: PR01434; NADHdGNASE5.
KM NAD: Oxidoreductase; Plastocyanone; Chloroplast.
FT NON_TER 686 686
SQ SEQUENCE 686 AA; 77797 MW; 28735E7A80FEA8CF CRC64;

Query Match 79.2%; Score 38; DB 8; Length 686;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 50 HOYVSWT 57

RESULT 12
ID 095645 PRELIMINARY; PRT; 688 AA.
AC 095645;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase (Fragment).
GN NDHF.
OS Gossypium hirsutum (Upland cotton).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Malvales; Malvaceae; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RA Seelanan T., Wendel J.F., Schnabel A.;
RT "Congruence and consensus in the cotton tribe: Evidence from the
RT nuclear and plastid genomes."
RL Syst. Bot. 0:0-0(1996).
DR EMBL: U55340; AAB17747.1; -
DR InterPro: IPR003916; NADHdb_oxred5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS: PR01434; NADHdGNASE5.
KM NAD: Oxidoreductase; Plastocyanone; Chloroplast.
FT NON_TER 1 1
SQ SEQUENCE 688 AA; 77857 MW; B1FE8BEA0A95DA57 CRC64;

Query Match 79.2%; Score 38; DB 8; Length 688;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 57 HOYVSWT 64

RESULT 13

OBHTM4
ID 08HTM4 PRELIMINARY; PRT; 701 AA.
AC 08HTM4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit F (Fragment).
GN NDHF.
OS Achlys triphylla (Vanilla leaf).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Berberidaceae; Achlys.
OX NCBI_TaxID=63345;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RA Kim Y.-D., Kim S.-H., Jansen R.K.;
RT "Phylogeny of the Berberidaceae based on the sequences of the
RT chloroplast gene ndhF."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY145158; AAN64401.1; -
KM Chloroplast.
FT NON_TER 701 701
SQ SEQUENCE 701 AA; 78560 MW; 7DB5860E7F53389A CRC64;

Query Match 79.2%; Score 38; DB 8; Length 701;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8
Db 69 HOYVSWT 76

RESULT 14
ID 09SC10 PRELIMINARY; PRT; 702 AA.
AC 09SC10;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit F (Fragment).
GN NDHF.
OS Tetracera asiatica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Dilleniaceae; Tetracera.
OX NCBI_TaxID=85285;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RA Albach D.C., Solits P.S., Solits D.E., Olmstead G.;
RT "Phylogeny of the Asteridae s.l. based on sequences from four
RT different genes."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ236277; CAB65467.1; -
DR InterPro: IPR003916; NADHdb_oxred5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR PRINTS: PR01434; NADHdGNASE5.
KM NAD: Oxidoreductase.
FT NON_TER 1 1
SQ SEQUENCE 702 AA; 78999 MW; ICB8A0F547FB76B2 CRC64;

Query Match 79.2%; Score 38; DB 10; Length 702;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HOYLSWT 8

DB 60 HQYVMSWT 67

RESULT 15

08HTM2 PRELIMINARY; PRT; 704 AA.
 AC 08HTM2; 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NADH dehydrogenase subunit F (Fragment).
 GN NDHF.
 OS Leontice evermannii.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 OC Berberidaceae; Leontice.
 OX NCBI_TaxID=211973;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim Y.-D., Kim S.-H., Jansen R.K.;
 RT "Phylogeny of the Berberidaceae based on the sequences of the
 RT chloroplast gene ndhF."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY145160; MAM64403.1; -.
 KW Chloroplast.
 FT NON_TER 704 704
 SQ SEQUENCE 704 AA; 78783 MW; 8EFEB8AEFB33ECA44 CRC64;

Query Match 79.2%; Score 38; DB 8; Length 704;
 Best Local Similarity 75.0%; Pred. NO. 86;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HQYVMSWT 8
 DB 69 HQYVMSWT 76

Search completed: October 7, 2003, 19:19:59
 Job time : 9.97241 secs

XX 12-AUG-1994; 94US-0289576.
PR (IMMU-) IMMUNOMEDICS INC.
XX
XX
XX Hansen H, Leung S;
PI WPI: 1996-139454/14.
XX N-PSDB; AAT13802.
DR
XX Chinese and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell lymphoma(s) and
PT leukaemia(s).
XX
XX Claim 5; Page 36-37; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). These were subsequently linked, respectively, to human
CC kappa and IgG1 constant regions. A humanised Mab was obtained that
CC retained the B-lymphoma and leukaemia cell targeting and
CC internalisation characteristics of the parental LL2 Mab, and which
CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
CC cytostatic agent for therapeutic appln.
XX
SQ Sequence 116 AA;
Query Match 100.0%; Score 620; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 6,1e-43;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QVQLDSGAELEKPGASVSKSCASGYTFSTYWLHWIKRPGGLEWIGYINPRNDYTEY 60
DB 1 QVQLDSGAELEKPGASVSKSCASGYTFSTYWLHWIKRPGGLEWIGYINPRNDYTEY 60
OY 61 NONFKDQATLTDKSSSTAYVMOLSLTSDSAVYYCARRDITTFYWGQGTTLTVSS 116
DB 61 NONFKDQATLTDKSSSTAYVMOLSLTSDSAVYYCARRDITTFYWGQGTTLTVSS 116
RESULT 2
AAM27696
ID AAM27696 standard; Protein; 116 AA.
XX
XX AAM27696;
AC
XX
XX 14-APR-1998 (first entry)
DT
XX
XX Variable heavy chain of Mab LL2.
DE
XX
XX Variable heavy chain; B cell; monoclonal antibody; Mab; LL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine;
KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
XX Mus sp.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 31..35
FT Region /note= "complementarity determining region 1"
FT Region 50..66 /note= "complementarity determining region 2"
FT Region 99..105 /note= "complementarity determining region 3"
FT Region /note= "complementarity determining region 3"
XX
XX W09734632-A1.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04196.

XX 20-MAR-1996; 96US-0013709.
PR (IMMU-) IMMUNOMEDICS INC.
XX
XX
XX Hansen H, Leung S, Qu Z;
PI WPI: 1996-479995/44.
XX N-PSDB; AAT98129.
DR
XX
XX Monoclonal antibody engineered to contain glycosylation site - in
PT non-FC constant heavy or light chain region, useful to diagnose or
PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
XX Example 3; Fig 4B; 88pp; English.
XX
XX The present sequence is the variable heavy chain of the
CC B cell specific monoclonal antibody (Mab) LL2, which contains an
CC engineered tripeptide N-glycan acceptor sequence. LL2 is a highly
CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
CC murine Mab. The Mab can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia. The glycosylation site allows a label or
CC therapeutic agent of increased size to be conjugated to the
CC carbohydrate moiety, without affecting the Mab's binding affinity
CC or specificity.
XX
SQ Sequence 116 AA;
Query Match 100.0%; Score 620; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 6,1e-43;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QVQLDSGAELEKPGASVSKSCASGYTFSTYWLHWIKRPGGLEWIGYINPRNDYTEY 60
DB 1 QVQLDSGAELEKPGASVSKSCASGYTFSTYWLHWIKRPGGLEWIGYINPRNDYTEY 60
OY 61 NONFKDQATLTDKSSSTAYVMOLSLTSDSAVYYCARRDITTFYWGQGTTLTVSS 116
DB 61 NONFKDQATLTDKSSSTAYVMOLSLTSDSAVYYCARRDITTFYWGQGTTLTVSS 116
RESULT 3
AAR92219
ID AAR92219 standard; Protein; 116 AA.
XX
XX AAR92219;
AC
XX
XX 28-MAY-1996 (first entry)
DT
XX
XX Humanised LL2 Mab VH region (version hLL2-1).
DE
XX
XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
KW leukaemia; therapy; diagnosis; complementarity determining region;
KW CDR; antibody engineering.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 31..35
FT Region /label= CDR1
FT Region 50..66 /label= CDR2
FT Region 99..105 /label= CDR3
XX
XX W09604925-A1.
XX
XX 22-FEB-1996.
XX
XX 11-AUG-1995; 95WO-US09641.
XX
XX 12-AUG-1994; 94US-0289576.

```
XX (IMM-) IMMUNOMEDICS INC.
PA Hansen H, Leung S;
XX WPI: 1996-139454/14.
DR
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell lymphoma(s) and
PT leukaemia(s).
XX
XX Example 1: Page 40; 70pp; English.
PS
XX The complementarily determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). In an alternative version a glutamine was introduced
CC at position 5 of the humanised VH (AAR92219) to include a PstI site
CC useful for subcloning. The humanised VK and VH were subsequently
CC linked, respectively, to human kappa and IgG1 constant regions. A
CC humanised Mab was obtd. that retained the B-lymphoma and leukaemia
CC cell targeting and internalisation characteristics of the parental
CC LL2 Mab, and which exhibited a lowered HAMA reaction. It can be
CC linked to a cytostatic agent for therapeutic appln.
XX
XX Sequence 116 AA;
SQ
Query Match 89.8%; Score 557; DB 17; Length 116;
Best Local Similarity 84.5%; Pred. No. 7.5e-38;
Matches 98; Conservative 14; Mismatches 4; Indels 0; Gaps 0;
QY 1 QVQLQESGAEISKPGASVKMSCKASGYFTSYWLHWIKRPGGLEWIGYINPRNDYTEY 60
DB 1 QVQLVQSGAEVKKRQSSVKSCAKSGYFTSYWLHWIRQAPGGGLEWIGYINPRNDYTEY 60
QY 61 NONPKDKATITLADSSSTAYMQSLSSSEDAVYYCARDDITTYWGQGITLTYS 116
DB 61 NONPKDKATITLADSTNTAYMELSLRSEDAFYFCARDITTYWGQGITVTYSS 116
RESULT 4
AAR92218
ID AAR92218 standard; Protein: 116 AA.
XX
XX AAR92218;
AC
XX
XX 28-MAY-1996 (first entry)
DE Humanised LL2 Mab VH region.
XX
XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
KW leukaemia; therapy; diagnosis; complementarity determining region;
KW CDR; antibody engineering.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Region 31..35
FT Region /label= CDR1
FT Region 50..66
FT Region /label= CDR2
FT Region 99..105
FT Region /label= CDR3
XX
XX MO9604925-A1.
XX
XX 22-FEB-1996.
XX
XX 11-AUG-1995; 95WO-US09641.
XX
XX 12-AUG-1994; 94US-0289576.
XX
```

```
PA (IMM-) IMMUNOMEDICS INC.
XX Hansen H, Leung S;
XX WPI: 1996-139454/14.
DR N-PSDB; NAT15804.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell lymphoma(s) and
PT leukaemia(s).
XX
XX Claim 5; Page 39; 70pp; English.
PS
XX The complementarily determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). These were subsequently linked, respectively, to human
CC kappa and IgG1 constant regions. A humanised Mab was obtd. that
CC retained the B-lymphoma and leukaemia cell targeting and
CC internalisation characteristics of the parental LL2 Mab, and which
CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
CC cytostatic agent for therapeutic appln.
XX
XX Sequence 116 AA;
SQ
Query Match 88.7%; Score 550; DB 17; Length 116;
Best Local Similarity 83.6%; Pred. No. 2.8e-37;
Matches 97; Conservative 14; Mismatches 5; Indels 0; Gaps 0;
QY 1 QVQLQESGAEISKPGASVKMSCKASGYFTSYWLHWIKRPGGLEWIGYINPRNDYTEY 60
DB 1 QVQLVQSGAEVKKRQSSVKSCAKSGYFTSYWLHWIRQAPGGGLEWIGYINPRNDYTEY 60
QY 61 NONPKDKATITLADSSSTAYMQSLSSSEDAVYYCARDDITTYWGQGITLTYS 116
DB 61 NONPKDKATITLADSTNTAYMELSLRSEDAFYFCARDITTYWGQGITVTYSS 116
RESULT 5
AAW27698
ID AAW27698 standard; Protein: 116 AA.
XX
XX AAW27698;
AC
XX
XX 14-APR-1998 (first entry)
DE Variable heavy chain of Mab hLL2.
XX
XX Variable heavy chain; B cell; monoclonal antibody; Mab; hLL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine; humanised;
KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
XX Chimeric - Mus sp.
OS
XX
XX Chimeric - Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 31..35
FT Region /note= "complementarity determining region 1"
FT Region 50..66
FT Region /note= "complementarity determining region 2"
FT Region 99..105
FT Region /note= "complementarity determining region 3"
XX
XX WO9734632-A1.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04196.
XX
XX 20-MAR-1996; 96US-0013709.
XX
```

PA (IMMU-) IMMUNOMEDICS INC.
 XX
 PI Hansen H, Leung S, Ou Z;
 DR WPI: 1997-479995/44.
 XX N-PSDB; AAT88131.
 XX
 PT Monoclonal antibody engineered to contain glycosylation site - in
 PT non-Fc constant heavy or light chain region, useful to diagnose or
 PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
 XX
 PS Example 3; Fig 5B; 88pp; English.
 XX
 CC The present sequence is the variable heavy chain of the
 CC B cell specific monoclonal antibody (mAb) hL2. hL2 is a highly
 CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
 CC huminised murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukaemia.
 XX
 SO Sequence 116 AA;
 50
 Query Match 88.7%; Score 550; DB 18; Length 116;
 Best Local Similarity 83.6%; Pred. No. 2.8e-37;
 Matches 97; Conservative 14; Mismatches 5; Indels 0; Gaps 0;
 QY 1 QVQLQSSGAELESPGASVYKMSKASGYTFTSYLHKIKRPGQGLEWIGYINPRNDYTEY 60
 DB 1 QVQLVSGAEVKKPGSSVSVKSCASGYTFTSYLHWVRQAPGGGLEWIGYINPRNDYTEY 60
 QY 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 DB 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 Db 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 RESULT 6
 AAR21279
 ID AAR21279 standard; Protein; 114 AA.
 AC AAR21279;
 XX
 DT 21-MAY-1992 (first entry)
 XX
 DE Murine VH group 1 chain P specific for phox.
 XX
 KW Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat;
 KW plus; g3p; binding; adsorption; gene VIII; diverse repertoire;
 KM specific binding pairs; replicable genetic display package.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Binding-site 31..35
 FT Binding-site /label= CDR1
 FT Binding-site 50..66
 FT Binding-site /label= CDR2
 FT Binding-site 99..103
 FT Binding-site /label= CDR3
 FT /note=" D/N-X-G-X-X motif "
 FT
 FT
 PN WO9201047-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 10-JUL-1991; 91WO-GB01134.
 XX
 XX 15-MAY-1991; 91GB-0010549.
 PR 10-JUL-1990; 90GB-0015198.
 PR 19-OCT-1990; 90GB-0022845.
 PR 12-NOV-1990; 90GB-0024503.
 PR 06-MAR-1991; 91GB-0004744.
 XX
 XX (CAMP-) CAMBRIDGE ANTIBODY.
 PA

PA (MED-) MED RES COUNCIL.
 XX
 PI McCafferty J, Pope AR, Johnson RS, Hoogenboom HRJ, Griffiths AD;
 PI Jackson RH, Holliger KP, Marks JD, Clackson TP, Chiswell DJ;
 PI Winter GP, Bonnett TP;
 XX
 DR WPI: 1992-056862/07.
 XX
 XX Producing members of specific binding pairs - by expression in
 PT recombinant host cells with a secreting replicable genetic
 PT display package.
 XX
 PS Example 22; Fig 24; 209pp; English.
 XX
 CC The VH sequence is one of 22 (AAR21264-85) expressed from a single
 CC chain Fv library. The library produces a diverse repertoire of
 CC antibody fragments specific for 2-phenyl-5-oxazolone (phox). It was
 CC prep. using cDNA generated from mRNA from mice immunised with phox
 CC coupled to chicken serum albumin. The VH and VL kappa sequences
 CC were separately amplified by PCR (AAQ23474-84) and ligated into
 CC fDCAV2 (AAQ23463) for expression on the phage surface as fusions with
 CC gene III. The resulting library of clones was diverse; 23 hapten
 CC binding clones were sequenced revealing 8 VH genes (A-H; AAR21264-71)
 CC in a variety of pairings with 7 Vk genes (a-g; AAR21286-92). Most
 CC clones were Vk-d combinations so a further hierarchical library was
 CC prep. by "crossing" Vk-d with the VH repertoire. The resulting
 CC library was screened for hapten binding and 24 clones sequenced. 13
 CC new partners (AAR21272-85) for Vk-d were identified. Nearly all the
 CC VH genes belonged to gp 1, with only one, "E", being of gp 2 (VHox1).
 CC Of the 24 hierarchical clones, only one was of type "P". The kd of
 CC VH-B/Vk-d for phox-GABA was 10 nM one of the highest values found.
 CC This suggests that phage bearing scfv fragments having weak affinities
 CC can be selected with antigen, probably due to the avidity of the
 CC multiple antibody heads on the phage. The different combinations
 CC could also be isolated on a basis of antigen affinity.
 CC See also AAR21260-307, 309-311; AAR2450, 565-581.
 CC
 CC
 CC
 CC
 SO Sequence 114 AA;
 50
 Query Match 86.1%; Score 534; DB 13; Length 114;
 Best Local Similarity 86.2%; Pred. No. 5.3e-36;
 Matches 100; Conservative 7; Mismatches 7; Indels 2; Gaps 1;
 QY 1 QVQLQSSGAELESPGASVYKMSKASGYTFTSYLHKIKRPGQGLEWIGYINPRNDYTEY 60
 DB 1 QVQLQSSGAELESPGASVYKMSKASGYTFTSYLHMVVRQAPGGGLEWIGYINPRNDYTEY 60
 QY 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 DB 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 Db 61 NQFKDKATLTADKSSSTAYMOLSLTSEDSAVYYCARRDITFTYGGGTTLVSS 116
 RESULT 7
 AAP93079
 ID AAP93079 standard; peptide; 159 AA.
 AC AAP93079;
 XX
 DT 25-MAR-2003 (updated)
 DT 31-OCT-2002 (updated)
 DT 15-MAR-1990 (first entry)
 XX
 DE Heavy chain of monoclonal antibody 6A4.
 XX
 KW Monoclonal antibody 6A4; heavy chain; Pseudomonas aeruginosa; OMP-1.
 XX
 OS Unidentified.
 XX
 PN EP338395-A.
 XX
 PD 25-OCT-1989.
 XX
 PF 12-APR-1989; 89EP-0106463.
 PA


```

XX 19-APR-1988; 88DE-3813023.
XX
XX (BEHW) BEHRINGERWERKE AG.
XX
XX Domdey H, Marget M, Vonspecht B;
XX WPI; 1989-310861/43.
XX DR N-PSDB; AAN91645.
XX
XX Monoclonal antibody to Pseudomonas aeruginosa - and DNA coding for
XX variable antibody regions.
XX
XX Claim 1; page 6; 7pp; german.
XX
XX The peptide is encoded by the heavy chain of monoclonal antibody 6A4.
XX 6A4 reacts with the OMP-1 protein of all 19 known serotypes of
XX P.aeruginosa. It is used for therapy and diagnosis of infection, and as
XX a carrier for drugs. The antibody is IgG2a subclass.
XX (Updated on 31-OCT-2002 to add missing OS field.)
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 25-MAR-2003 to correct PI field.)
XX
XX
XX Sequence 159 AA;
SQ
Query Match 85.3%; Score 529; DB 10; Length 159;
Best Local Similarity 82.5%; Pred. No. 1.9e-35;
Matches 99; Conservative 8; Mismatches 9; Indels 4; Gaps 1;
QY 1 QVOLOESGAELESKRGASVYKMSCKASGYTFYTHMTIKORPGGLEWIGYINPNDYTEY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 20 QVOLOQSGAELEADPGASVYKMSCKASGYTFYTHMTIKORPGGLEWIGYINPNDYTEY 79
QY 61 NONFKDKATLTADKSSSTAYVQSLTSEDSAVYYCAR---DITFFYMGQGTLLTVSS 116
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 80 NQNRQDKATLTADKSSSTAYVQSLTSEDSAVYYCARWGNGSITGLDYWGCGTSTVYSS 139
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
AAW76127
ID AAW76127 standard; Protein; 138 AA.
XX
XX AAW76127;
AC
XX
XX 20-NOV-1998 (first entry)
DT
XX
XX Murine ICR-1.1 V-H region PCR product protein.
DE
XX
XX Interleukin adhesion molecule; ICM-R; modulator; 14.3.3 family;
KW HSI-beta; tubulin; inhibitor; stimulator; effector; immune response;
KW inflammation; disorder; T cell activation; macrophage; Crohn's disease;
KW adult respiratory distress syndrome; stroke; multiple sclerosis; asthma;
KW rheumatoid arthritis; tumour growth; human immune deficiency virus;
KW infection; diabetes; graft vs. host disease; passive immunisation.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 2
FT Misc-difference /label- unknown
FT Misc-difference 6 /label- unknown
FT Misc-difference 7 /label- unknown
FT Misc-difference /label- unknown
FT Misc-difference 8 /label- unknown
FT
XX US5773218-A.
PN
XX 30-JUN-1998.
PD
XX
XX 07-JUN-1995; 95US-0482882.
PF
XX

```

```

PR 05-AUG-1994; 94US-0286754.
PR 27-JAN-1992; 92US-0827689.
PR 26-MAY-1992; 92US-0889724.
PR 05-JUN-1992; 92US-0894061.
PR 22-JAN-1993; 93US-0009286.
PR 26-JAN-1993; 93WO-0500787.
PR 05-AUG-1993; 93US-0102852.
PR 07-JUN-1995; 95US-0482882.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX WPI; 1998-386989/33.
XX DR N-PSDB; AAV56413.
XX
XX Identifying compounds that modulate interaction of intercellular
XX adhesion molecule R - with ligands HSI-beta and tubulin using
XX two-hybrid assay, useful for treating inflammation, T cell
XX activation etc.
XX
XX Example 13; Column 125-128; 108pp; English.
XX
XX This sequence represents a murine ICR-1.1 V-H region amplified PCR
XX product. This sequence is used in the isolation of a novel human
XX intercellular adhesion molecule, ICM-R. This sequence is used in a
XX method which investigates modulators of the interaction between ICM-R
XX and the 14.3.3 family member HSI-beta and tubulin. An anti-ICM-R
XX peptide, can block, inhibit or stimulate ligand/receptor interactions
XX involving ICM-R, particularly its effector functions involved in
XX (non)specific immune responses. ICM-R related agents may be used to
XX treat or monitor inflammation, disorders involving T cell activation or
XX macrophages, e.g. adult respiratory distress syndrome, stroke, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis, asthma, tumour growth,
XX human immune deficiency virus infection, diabetes, graft vs. host disease
XX and many others. Antibodies may also be used for passive immunisation,
XX for purifying, detecting or quantifying ICM-R and for identifying
XX ICM-R expressing cells.
XX
XX Sequence 138 AA;
SQ
Query Match 84.6%; Score 524.5; DB 19; Length 138;
Best Local Similarity 82.4%; Pred. No. 3.8e-35;
Matches 98; Conservative 9; Mismatches 9; Indels 3; Gaps 1;
QY 1 QVOLOESGAELESKRGASVYKMSCKASGYTFYTHMTIKORPGGLEWIGYINPNDYTEY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 20 QVOLOQSGAELEADPGASVYKMSCKASGYTFYTHMTIKORPGGLEWIGYINPNDYTEY 79
QY 61 NONFKDKATLTADKSSSTAYVQSLTSEDSAVYYCAR---DITFFYMGQGTLLTVSS 116
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 80 NQNRQDKATLTADKSSSTAYVQSLTSEDSAVYYCARWGNGSITGLDYWGCGTSTVYSS 138
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
AAW71257
ID AAW71257 standard; Protein; 138 AA.
XX
XX AAW71257;
AC
XX
XX 25-MAR-2003 (updated)
DT 18-NOV-1998 (first entry)
DT
XX
XX Murine antibody ICR-1.1 heavy chain amino acid sequence.
DE
XX
XX Human; ICM-R; intercellular adhesion molecule; adhesion; treatment;
KW inflammatory condition; asthma; tumour growth; metastasis;
KW viral infection; antibody ICR-1.1.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH
XX

```

```

FT Misc-difference 2 /note- "encoded by GRA"
FT Misc-difference 6 /note- "encoded by RTC"
FT Misc-difference 7 /note- "encoded by WTB"
FT Misc-difference 8 /note- "encoded by HTC"
XX
XX US5811517-A.
XX
XX 22-SEP-1998.
XX
XX
XX 07-JUN-1995; 95US-0483389.
XX
XX 05-AUG-1994; 94US-0286754.
XX 26-JAN-1993; 93WO-US00787.
XX 27-JAN-1992; 92US-0827689.
XX 26-MAY-1992; 92US-0889724.
XX 05-JUN-1992; 92US-0894061.
XX 22-JAN-1993; 93US-0009266.
XX 03-AUG-1993; 93US-0102852.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX WPI; 1998-530940/45.
XX N-PSDB; AAV54863.
XX
XX DNA encoding mutant ICAM-R polypeptide(s) - useful for diagnosis
XX and treatment of cell adhesion based disease conditions e.g.
XX inflammation or asthma
XX
XX Example 13; Columns 125-126; 11pp; English.
XX
XX The present sequence represents the heavy chain of murine antibody
XX ICR-1.1. This antibody is specific for ICAM-R (intercellular adhesion
XX molecule-R). ICAMs are polypeptides that are expressed on blood vessel
XX endothelial cell surfaces and are involved in the adhesion events in
XX various conditions. ICAM-R variants (see AAW1264-69) can be used to
XX treat or monitor inflammatory conditions involving specific or
XX non-specific immune responses, asthma, tumour growth and/or metastasis
XX and viral infections. The ICAM variants are produced recombinantly, from
XX expression libraries of mutated sequences, and the ones that are
XX claimed are the ones that have been found to be especially involved in
XX adhesion events. They can also be used to raise antibodies, also for
XX use as therapeutic or diagnostic agents.
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 138 AA:
SQ
Query Match 84.6%; Score 524.5; DB 19; Length 138;
Best Local Similarity 82.4%; Pred. No. 3.8e-35;
Matches 98; Conservative 9; Mismatches 9; Indels 3; Gaps 1;
QY 1 QVQLDSGAEISPGASVSKASGYTFTSWLMHKRPGQGLEWIGYINPRNDYTEY 60
DB 20 QVQLDSGAEISPGASVSKASGYTFTSWLMHKRPGQGLEWIGYINPRNDYTEY 79
QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYVCARDITFT--YWGQGTTLTVSS 116
DB 80 NORFDKATLTADKSSSTAYMQLSLTSDSAVYVCARDIGNSYGLDWGQGTSTVSS 138
RESULT 10
AAV00783
ID AAY00783 standard; Protein; 138 AA.
XX
XX AAY00783;
AC AAY00783;
XX
XX 14-MAY-1999 (first entry)
XX

```

```

DE Antibody against ICAM-R.
XX
XX ICAM; immunoglobulin-like loop; intercellular adhesion molecule receptor;
XX alpha d/CD18; antibody; immunisation; inflammatory response; asthma;
XX tumour growth; viral infection; therapy.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..138 /note- "Xaa- unspecified amino acid"
XX
XX US5880268-A.
XX
XX 09-MAR-1999.
XX
XX 07-JUN-1995; 95US-0483932.
XX
XX 05-AUG-1994; 94US-0286754.
XX 27-JAN-1992; 92US-0827689.
XX 26-MAY-1992; 92US-0889724.
XX 05-JUN-1992; 92US-0894061.
XX 22-JAN-1993; 93US-0009266.
XX 26-JAN-1993; 93WO-US00787.
XX 05-AUG-1993; 93US-0102852.
XX 07-JUN-1993; 93US-0483932.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX WPI; 1999-204041/17.
XX N-PSDB; AAX21879.
XX
XX New intercellular adhesion molecule receptor (ICAM-R) specific
XX antibodies - useful for modulating ligand/receptor binding and
XX biological activities involving ICAM-R, especially those of the
XX specific and non-specific immune systems
XX
XX Example 13; Column 125-126; 108pp; English.
XX
XX This sequence represents an antibody specific for ICAM-R.
XX The invention relates to antibodies (Ab) which bind specifically
XX to the intercellular adhesion molecule receptor (ICAM-R), inhibiting the
XX interaction between ICAM-R and alpha d/CD18. Abs with specific ICAM-R
XX binding are useful in compositions for immunisation, and for purifying
XX ICAM-R polypeptides and identifying cells expressing ICAM-R on their cell
XX surface, modulating ligand/receptor binding and biological activities
XX involving ICAM-R, especially inflammatory responses of the specific
XX immune system, the non-specific immune system, monitoring and treating
XX asthma, tumour growth, and/or metastasis, and viral infection (e.g. HIV
XX infection). In particular diseases involving an essential T cell
XX activation (e.g. asthma, psoriasis, diabetes, graft vs. host disease,
XX tissue transplant rejection, and multiple sclerosis) may be treated with
XX anti-ICAM-R antibodies. The Abs specifically bind to and identify ICAM-R
XX and disrupt ICAM-R to cell adhesion molecule, especially alpha d/CD18
XX binding.
XX
XX Sequence 138 AA:
SQ
Query Match 84.6%; Score 524.5; DB 20; Length 138;
Best Local Similarity 82.4%; Pred. No. 3.8e-35;
Matches 98; Conservative 9; Mismatches 9; Indels 3; Gaps 1;
QY 1 QVQLDSGAEISPGASVSKASGYTFTSWLMHKRPGQGLEWIGYINPRNDYTEY 60
DB 20 QVQLDSGAEISPGASVSKASGYTFTSWLMHKRPGQGLEWIGYINPRNDYTEY 79
QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYVCARDITFT--YWGQGTTLTVSS 116
DB 80 NORFDKATLTADKSSSTAYMQLSLTSDSAVYVCARDIGNSYGLDWGQGTSTVSS 138

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 60 Seconds
(without alignments)
305,878 Million cell updates/sec

Title: US-09-988-013a-4
Perfect score: 620
Sequence: 1 QVOLOESGAELEKSGASVKM.....ARRDITFYWGQGTTLVSS 116

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620	100.0	116	10	US-09-741-843-4
2	620	100.0	116	11	US-09-894-839-4
3	620	100.0	116	11	US-09-988-013a-4
4	557	89.8	116	10	US-09-741-843-9
5	557	89.8	116	11	US-09-894-839-21
6	557	89.8	116	11	US-09-988-013a-9
7	550	88.7	116	10	US-09-741-843-8
8	550	88.7	116	11	US-09-894-839-8
9	550	88.7	116	11	US-09-988-013a-8
10	524.5	84.6	138	9	US-09-753-436-78
11	523	84.4	116	12	US-10-127-890-167
12	523	84.4	135	12	US-10-244-821-91
13	519	83.7	438	12	US-10-244-821-88
14	507	81.8	331	15	US-10-059-261-169
15	505.5	81.3	119	11	US-09-795-515-30

16	505.5	81.5	119	15	US-10-267-286A-10	Sequence 10, Appl
17	505.5	81.5	468	11	US-09-795-515-7	Sequence 7, Appl1
18	501	80.8	116	12	US-10-127-890-168	Sequence 168, App
19	499	80.5	256	15	US-10-247-488-2	Sequence 2, Appl1
20	499	80.5	258	15	US-10-247-488-4	Sequence 4, Appl1
21	496	80.0	120	15	US-10-096-246-12	Sequence 12, Appl1
22	491.5	79.3	500	12	US-10-168-809-22	Sequence 22, Appl
23	484	78.1	242	16	US-10-259-087A-18	Sequence 18, Appl
24	483	77.9	127	10	US-09-988-831-7	Sequence 7, Appl1
25	483	77.9	127	12	US-10-373-561-7	Sequence 7, Appl1
26	482.5	77.8	143	8	US-08-779-784-26	Sequence 26, Appl
27	482	77.7	122	15	US-10-096-246-10	Sequence 10, Appl
28	482	77.7	122	15	US-10-096-246-11	Sequence 11, Appl
29	480	77.4	242	16	US-10-259-087A-20	Sequence 20, Appl
30	478.5	77.2	119	15	US-10-233-996-39	Sequence 39, Appl
31	477.5	77.0	140	15	US-10-283-349-27	Sequence 27, Appl
32	476	76.8	122	8	US-08-779-784-28	Sequence 28, Appl
33	473.5	76.4	116	11	US-09-940-727B-17	Sequence 17, Appl
34	473.5	76.4	464	12	US-10-384-933-9	Sequence 9, Appl1
35	473.5	76.4	464	15	US-10-216-484-9	Sequence 9, Appl1
36	471	76.0	113	11	US-09-940-727B-118	Sequence 118, App
37	469.5	75.7	136	15	US-10-138-505-12	Sequence 12, Appl
38	469.5	75.7	245	15	US-10-138-505-40	Sequence 40, Appl
39	469.5	75.7	271	15	US-10-138-505-34	Sequence 34, Appl
40	469.5	75.7	274	15	US-10-138-505-32	Sequence 32, Appl
41	469	75.6	116	12	US-10-127-890-169	Sequence 169, App
42	464.5	74.9	123	9	US-09-753-436-86	Sequence 86, Appl
43	464.5	74.9	412	12	US-10-244-821-6	Sequence 6, Appl1
44	464.5	74.9	412	15	US-10-013-173-6	Sequence 6, Appl1
45	464.5	74.9	412	15	US-10-150-762-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-741-843-4
Sequence 4, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
FILE REFERENCE: 018733/0996
CURRENT FILING DATE: 2000-12-22
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 116
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-741-843-4
Query Match 100.0%; Score 620; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.1e-49;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVOLOESGAELEKSGASVKMCKASGTYFTYMLHWIKORPGGLEWYGYINPRNDYEV 60
|||||
DB 1 QVOLOESGAELEKSGASVKMCKASGTYFTYMLHWIKORPGGLEWYGYINPRNDYEV 60
|||||
QY 61 NONFKDKATLADSSSTAYWQLSLTSEDAVYYCARDDITFYWGQGTTLVSS 116
|||||
DB 61 NONFKDKATLADSSSTAYWQLSLTSEDAVYYCARDDITFYWGQGTTLVSS 116
|||||

```

RESULT 2
US-09-894-839-4
; Sequence 4, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-4

Query Match
Best Local Similarity 100.0%; Score 620; DB 11; Length 116;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60
DB 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60

QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116
DB 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116

RESULT 3
US-09-988-013A-4
; Sequence 4, Application US/0988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-988-013A-4

Query Match
Best Local Similarity 100.0%; Score 620; DB 11; Length 116;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60
DB 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60

QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116
DB 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116

```

```

QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116
DB 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116

RESULT 4
US-09-741-843-9
; Sequence 9, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: AND LEUKEMIA CELLS
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-9

Query Match
Best Local Similarity 89.8%; Score 557; DB 10; Length 116;
Matches 98; Conservative 14; Mismatches 4; Indels 0; Gaps 0;

QY 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60
DB 1 OVOLOESGAELEKPGASVMSCKASGYTFTSYLHWIKORPGGLEWIGYINPRNDYTEY 60

QY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116
DB 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYYCARDITTFYWGQTTLVSS 116

RESULT 5
US-09-894-839-21
; Sequence 21, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-21

Query Match
Best Local Similarity 89.8%; Score 557; DB 11; Length 116;
Matches 98; Conservative 14; Mismatches 4; Indels 0; Gaps 0;

```


[illegible]

```

; RESULT 6
; US-09-988-013A-9
; Sequence 9, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Shul-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B
; TITLE OF INVENTION: CELL LYMPHOMA AND LEUKEMIA CELLS
; FILE REFERENCE: 18735/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-988-013A-9

```

```

1      RESULT 7
2      US-09-741-843-8
3      : Sequence 8, Application US/09741843
4      : Patent No. US20020102254A1
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: LEUNG, Shui-on
9      : APPLICANT: HANSEN, Hans
10     :
11     : TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA
12     :
13     : TITLE OF INVENTION: AND LEUKEMIA CELLS
14     :
15     : FILE REFERENCE: 018733/0996
16     :
17     : CURRENT APPLICATION NUMBER: US/09/741,843
18     :
19     : CURRENT FILING DATE: 2000-12-22
20     :
21     : PRIOR APPLICATION NUMBER: US 09/127,902
22     :
23     : PRIOR FILING DATE: 1998-08-03
24     :
25     : PRIOR APPLICATION NUMBER: US 08/690,102
26     :
27     : PRIOR FILING DATE: 1996-07-06
28     :
29     : PRIOR APPLICATION NUMBER: US 08/289,576
30     :
31     : PRIOR FILING DATE: 1994-08-12
32     :
33     : NUMBER OF SEQ ID NOS: 21
34     :
35     : SOFTWARE: patentin version 3.1
36     :
37     : SEQ ID NO 8
38     :
39     : LENGTH: 116
40     :
41     : TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-09-741-843-8

Query Match      88.7%; Score 550; DB 10; Length 116;
Best Local Similarity 83.6%; Pred. No. 2,9e+43;
Matches 97; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 1 QVQLQESGAEELSKPGASVKMSCKASGYTFSTYWLMIKQRPQGGLEWIGYINPRNDYREY 60
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1 QVQLVQSGAEAEKRRKGGSSVKSCAKSGYFTSTYWLHWNAQAGGGLIEWIGYINPRNDYREY 60

QY 61 NQNKDKATLTADKSSSTAYYMQLSLSEDSAAVYYCARNDITTFYWGQITLVSS 116
   ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 61 NQNKDKATLTADKSSSTAYYMQLSLSEDSIAFAFCARNDITTFYWGQITVTVSS 116

```

```

RESULT 8
US-09-894-839-8
: Sequence 8, Application US/09894839
: Publication NO. US20030035800A1
: GENERAL INFORMATION:
: APPLICANT: LEIDING, Shui-on
: APPLICANT: HANSEN, Hans
: APPLICANT: QU, Zhengxing
: TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
: FILE REFERENCE: 018733/1049
: CURRENT APPLICATION NUMBER: US/09/894, 839
: PRIORITY FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 09/155,107
: PRIORITY FILING DATE: 1998-11-17
: PRIOR APPLICATION NUMBER: US 20/013,709
: PRIORITY FILING DATE: 1996-03-20
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 8
: LENGTH: 116
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-894-839-8

```

RESULT 9
US-09-988-013A-8
: Sequence 8, Application US/09988013A
: Publication No. US20030103979A1
: GENERAL INFORMATION:
: APPLICANT: LEUNG, Shui-on
: APPLICANT: HANSEN, Hans
: TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
: TITLE OF INVENTION: CELL LYMPHOMA AND LEUKEMIA CELLS
: FILE REFERENCE: 18733/1082
: CURRENT APPLICATION NUMBER: US/09/988, 013A
: CURRENT FILING DATE: 2002-10-29
: PRIOR APPLICATION NUMBER: US 09/741, 843
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/127, 902
: PRIOR FILING DATE: 1998-08-03
: PRIOR APPLICATION NUMBER: US 08/690, 102
: PRIOR FILING DATE: 1996-07-06
: PRIOR APPLICATION NUMBER: US 08/289, 576
: PRIOR FILING DATE: 1994-08-12

NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-988-013a-8

Query Match 88.7%; Score 550; DB 11; Length 116;
Best Local Similarity 83.6%; Pred. No. 2.9e-43;
Matches 97; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 1 OYOLGSGAELSPGASVMSCKASGYFTSYLHWIKRPGGLEWIGYINPRNDYTY 60
DB 1 OYOLGSGAELSPGASVMSCKASGYFTSYLHWIKRPGGLEWIGYINPRNDYTY 60

QY 61 NONFKDKATLTADKSSSTAYMOLSLTSDSAVYYCCARDITFTFYWGQTTLVSS 116
DB 61 NONFKDKATLTADKSSSTAYMOLSLTSDSAVYYCCARDITFTFYWGQTTLVSS 116

RESULT 10
US-09-753-436-78
Sequence 78, Application US/09753436
Patent No. US20010029293A1
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemay
TITLE OF INVENTION: ICAM-Related Materials and Methods
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/753,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/382,289
FILING DATE:
APPLICATION NUMBER: US 08/487,113
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,754
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Joseph A., Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 33282

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-753-436-78

Query Match 84.6%; Score 524.5; DB 9; Length 138;
Best Local Similarity 82.4%; Pred. No. 7.5e-41;
Matches 98; Conservative 9; Mismatches 9; Indels 3; Gaps 1;

QY 1 OYOLGSGAELSPGASVMSCKASGYFTSYLHWIKRPGGLEWIGYINPRNDYTY 60
DB 20 OYOLGSGAELADPGASVMSCKASGYFTSYLHWIKRPGGLEWIGYINPRNDYTY 79

QY 61 NONFKDKATLTADKSSSTAYMOLSLTSDSAVYYCCARDITFTFYWGQTTLVSS 116
DB 80 NONFKDKATLTADKSSSTAYMOLSLTSDSAVYYCCARDITFTFYWGQTTLVSS 138

RESULT 11
US-10-127-890-167
Sequence 167, Application US/10127890
Publication No. US2003016196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70, P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248

```
; INFORMATION FOR SEQ ID NO: 167:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 116 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 167:
US-10-127-890-167

Query Match      84.4%; Score 523; DB 12; Length 116;
Best Local Similarity 86.2%; Pred. No. 8.7e-41;
Matches 100; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 OVOLOESGAELSKPGASVKMSCKASGYTFSTYMLHWIKORPGGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 OVOLOOSGAELAKGASVKMSCKASGYTFSTYRMHWKORPGGLEWIGYINPSTGYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY 61 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARDDITTFYWGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARGGGVFDYWGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
US-10-244-821-91
; Sequence 91, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-244-821-91

Query Match      84.4%; Score 523; DB 12; Length 135;
Best Local Similarity 86.2%; Pred. No. 1e-40;
Matches 100; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 OVOLOESGAELSKPGASVKMSCKASGYTFSTYMLHWIKORPGGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 20 OVOLOOSGAELAKGASVKMSCKASGYTFSTYRMHWKORPGGLEWIGYINPSTGYTEY 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY 61 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARDDITTFYWGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 80 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARGGGVFDYWGQGTTLTVSS 135
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
US-10-244-821-88
; Sequence 88, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
```

```
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-244-821-88

Query Match      83.7%; Score 519; DB 12; Length 438;
Best Local Similarity 86.1%; Pred. No. 8e-40;
Matches 99; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 OVOLOESGAELSKPGASVKMSCKASGYTFSTYMLHWIKORPGGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 23 OVOLOOSGAELAKGASVKMSCKASGYTFSTYRMHWKORPGGLEWIGYINPSTGYTEY 82
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY 61 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARDDITTFYWGQGTTLTVSS 115
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 83 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARGGGVFDYWGQGTTLTVSS 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
US-10-059-261-169
; Sequence 169, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; TITLE OF INVENTION: (PTPC)
; FILE REFERENCE: 03495.0216
; CURRENT APPLICATION NUMBER: US/10/059,261
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 169
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector pAcGp67-ScFv350
US-10-059-261-169

Query Match      81.8%; Score 507; DB 15; Length 331;
Best Local Similarity 82.2%; Pred. No. 7.5e-39;
Matches 97; Conservative 9; Mismatches 10; Indels 2; Gaps 2;

OY 1 OVOLOESGAELSKPGASVKMSCKASGYTFSTYMLHWIKORPGGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 86 OVOLOOSGAELAKGASVKMSCKASGYTFSTYRMHWKORPGGLEWIGYINLSSGYTRY 145
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY 61 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCAR-RDITTF-YWGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 146 NONFKDKATLTADKSSSTAYWQLSLTSEDSAVYYCARAAQATTFDYWGQGTTLTVSS 203
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
US-09-795-515-30
; Sequence 30, Application US/09795515
; Publication No. US20030039645A1
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 43.2 Seconds

(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013a-4

Perfect score: 620

Sequence: 1 QVQLQESGAEISKPGASVKM.....ARRDITFFYWGCGTTLTVSS 116

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	85.3	139	2	Ig heavy chain pre
2	511.5	82.5	117	2	PL7-6 antibody hea
3	496	80.0	120	2	Ig heavy chain V r
4	494.5	79.8	120	2	Ig heavy chain V r
5	487	78.5	116	2	antibody fab Jcl 1
6	487	78.5	141	2	Ig heavy chain pre
7	486	78.4	120	2	Ig heavy chain V r
8	485	78.4	131	2	Ig heavy chain V r
9	482.5	77.8	123	2	Ig heavy chain V r
10	482	77.7	139	1	Ig heavy chain pre
11	482	77.7	287	4	PC4402
12	480.5	77.5	122	2	Ig heavy chain V r
13	478	77.1	118	2	Ig heavy chain V r
14	476	76.8	106	2	Ig heavy chain V r
15	475	76.6	118	2	Ig heavy chain V r
16	475	76.6	126	2	Ig heavy chain V r
17	475	76.6	136	2	Ig heavy chain pre
18	474.5	76.5	131	2	Ig heavy chain pre
19	474	76.5	139	2	Ig heavy chain pre
20	473.5	76.4	116	2	Ig heavy chain pre
21	473.5	76.4	118	2	Ig heavy chain V r
22	472.5	76.2	119	2	Ig heavy chain V r
23	470	75.8	122	2	Ig heavy chain V r
24	469.5	75.8	125	2	Ig heavy chain V r
25	469	75.6	114	2	Ig heavy chain V r
26	468.5	75.6	138	2	Ig heavy chain V r
27	468	75.5	117	2	Ig heavy chain V r
28	467.5	75.4	136	2	Ig heavy chain V r
29	467.5	75.4	140	2	Ig heavy chain V r

30	466.5	75.2	138	2	E32513	Ig heavy chain pre
31	464	74.8	115	2	C27563	Ig heavy chain V r
32	463.5	74.8	119	2	PL0089	Ig heavy chain V r
33	461	74.4	115	2	A54378	Ig heavy chain V r
34	461	74.4	141	2	A39276	Ig heavy chain pre
35	460	74.2	96	2	S17616	Ig heavy chain V r
36	460	74.2	120	2	G28195	Ig heavy chain V r
37	459	74.0	112	2	A30502	Ig heavy chain V r
38	458.5	74.0	117	2	G45722	Ig heavy chain V r
39	458	73.9	135	2	A30577	Ig heavy chain pre
40	458	73.9	144	2	B30502	Ig heavy chain V r
41	457	73.7	474	1	G2MS11	Ig gamma-2b chain
42	456	73.5	123	2	S60067	Ig heavy chain V r
43	456	73.5	128	2	I37267	Ig heavy chain V r
44	456	73.5	137	1	G2MS43	Ig heavy chain pre
45	455.5	73.5	140	1	HVMSG7	Ig heavy chain pre.

ALIGNMENTS

RESULT 1

PS0024 Ig heavy chain precursor V region (6A4) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 22-Nov-1996

C/Accession: PS0024

R/Margot, M.; Echardt, A.; Ehret, W.; von Specht, B.U.; Duchene, M.; Dondey, H.

A/Title: Cloning and characterization of cDNAs coding for the heavy and light chains

A/Reference number: PS0023; MUID:89232725; PMID:3149944

A/Accession: PS0024

A/Molecule type: mRNA

A/Residues: 1-139 <MAR>

A/Experimental source: strain BALB/c

C/Comment: This chain is obtained from an IgG2a monoclonal antibody against Pseudomon

C/Superfamily: Immunoglobulin V region; Immunoglobulin homology

C/Keywords: heterotrimer; signal sequence; status predicted <Sig>

F:1-19/Domain: signal sequence #status predicted <Sig>

F:20-139/Domain: Ig heavy chain V region #status predicted <IGV>

F:34-117/Domain: Immunoglobulin homology <IMM>

F:20/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic

Query Match 85.3%; Score 529; DB 2; Length 139;

Best local similarity 82.5%; Pred. No. 2e-40;

Matches 99; Conservative 8; Mismatches 9; Indels 4; Gaps 1;

QY 1 QVQLQESGAEISKPGASVKSGTFTSYLHMTKORPGGLEWIGTINRNDYTEY 60

Db 20 QVQLQESGAEISKPGASVKSGTFTSYLHMTKORPGGLEWIGTINRNDYTEY 79

QY 61 NONFKDKATLTADKSSSTAYMOLSLTSEDSAYYYCAR---DITFFYWGCGTTLTVSS 116

Db 80 NONFKDKATLTADKSSSTAYMOLSLTSEDSAYYYCTRSYNTYEGANDYWGCGTTLTVSS 139

RESULT 2

JC2269

PL7-6 antibody heavy chain - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 07-May-1999

C/Accession: JC2269; PC2186

R/Kurume, T.; Katayama, M.; Murakami, K.; Hashino, K.; Kamihagi, K.; Yasunoto, M.; Ka

J. Blochem. 115, 608-614, 1994

A/Title: Expression of recombinant mouse/human chimeric antibody specific to human GM

A/Reference number: JC2269; MUID:94334310; PMID:7520038

A/Accession: JC2269

A/Molecule type: mRNA

A/Residues: 1-117 <KUR>

A/Accession: PC2186

A/Molecule type: protein

A/Residues: 2-27 <KUR>

A/Experimental source: hybridoma cell

C:Comment: This protein is specific to human P-selectin.
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:2-98/Region: V segment
F:15-98/Domain: immunoglobulin homology <IM>
F:99-106/Region: D segment
F:107-117/Region: J segment

Query Match	82.5%	Score 511.5;	DB 2;	Length 117;
Best Local Similarity	82.1%	Pred. No. 6.2e-39;		
Matches 96; Conservative	9;	Mismatches 11;	Indels 1;	Gaps 1

[illegible]

```

61 NQNFQDKATLTADKSSSTAYMQLSSLSSEDSANVYYCARDDITTF-YWGCGTTLTVSS 116
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NQNFQDKATLTADKSSSTAYMQLSSLSSEDSANVYYCAGNPANFAYWGGTLYTVSA 117

```

RESULT 3
S41394

C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C:Accession: S41394
R:Margaritte, C.; Gilbert, D.; Brard, F.; Tiron, F.
R:Mus musculus (house mouse)
R:Mus musculus (house mouse)

A:Description: Structural characterization of an (NZB X NZW)F1 mouse-derived Igm anti-DNA antibody
A:Reference number: S41393
A:Accession: S41394
A:Creation: 1994-01-01
A:Release: 1994-01-01

A:Cross-references: EMBL:z29586; NID:g452354; PDB:CAA0703.1; PID:q13340808
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match	80.0%;	Score 496;	DB 2;	Length 120;
Best Local Similarity	77.5%;	Pred. No. 1.5e-37;		
Matches 93;	Conservative 11;	Mismatches 12;	Indels 4;	Gaps 1

Qy	1	OYOLOESAEELKSPASVKMSCKASGYFTSWLWIMKORPOGLEMIGYIRPADRYEX	60
		: : : : : : : : : :	
Db	1	EVOLQOASAEELVKRPAASVKLSCSKASGYFTSWLWIMWVQORPOGLEMIGELIDPSOYTY	60
Qy	61	NONRKDKATLTADKSSSTAYIMQSLSSIEDSDAVYICARR---DITTFYWGQTLVLNYS	116
		: : : : : : : : :	
Db	61	NONRKGRKTLTVNDSSSTAYIMQFSSLTSEDSDAVYICARRYGSRVSMDSMDYGGISTVWYS	120

RESULT 4
S25175

Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C:date: 20-Feb-1995 #sequence_rev:10n 20-Feb-1995 #text_change 23-Jul-1995
C:Accession: S25175
E:Monestier, M.; Fasy, T.M.; Losman, M.J.; Novick, K.E.; Muller, S.

A1:Description: Structure and binding properties of monoclonal antibodies to core histone
A1:Reference number: S25174

A: Molecule type: mRNA
A: Residues: 1-120 <MON>

A:Cross-references: EMBL:X67620; NID:g51056; PIDD:CAA7878.1; PID:g938266
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
E:15-98/Domain: immunoglobulin homology <IIM>

Query Match	79.88;	Score 494.5;	DB 2;	Length 120;
Best Local Similarity	78.38;	Pred. NO. 2.1e-37;		

Matches 94; Conservative 8; Mismatches 13; Indels 5; Gaps 1;

QY I QVQLDSEGAELSKPGASYKMSCASGYYFTSYMLHWIKRPRGGGLWIGIYNPRNDYTEY 600
| | | | : | | | | | | | | | | | | | | : |
Db I QVQLDQPGAEIVKPGASVKLSCASGYFTSYMMHWIKRPRGGGLWIGINDPSDETHY 600

```

61 NQKKDKATLTADKSSSTAYMQLSSITSEDSAVVYCARDDTT-----FYMGQGTLLVS 115
    || ||||| |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
db 61 NQKKDKATLTVDKSSNTAYMQLSSITSEDSAVFYCARKEITDDYNYFDYMGQGTLLVS 120

```

RESULT 5
S53751.

antibody Fab JEL 103 heavy chain - mouse
C/Species: Mus musculus (house mouse)
C/Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999
C/Accession: S53751
R/Pokkurati, P.R.; Bouthillier, F.; Li, Y.; Kuderova, A.; Lee, J.; Cygler, M.
J. Mol. Biol. 243, 283-297, 1994
A/Title: Preparation, characterization and crystallization of an antibody Fab
A/Reference number: S53750; MUID:95018269; PMID:753684

C; Superfamily: Immunoglobulin V region; Immunoglobulin homology
F; 15-98/Domain: Immunoglobulin homology <IMM>

Query Match	78.5%	Score 487;	DB 2;	Length 116;
Best Local Similarity	80.2%;	Pred. No. 9.4e-37;		
Matches 93;	Conservative 6;	Mismatches 17;	Indels 0;	Gaps 0;

Qy 1 QVQLDSGAELTSPGASVKMSCASGYFTSTYWLHM IKQRPGGLEWIGYINPRANDYTEY 600
|||:|||:||||:||||:|:|||||:|:|
Db 1 QVQLDQSGAELYPGASVKLTSCASGYFTSYNQMWKQRPGGLEWIGELDPSDYTNV 600

QY 61 NQNFKDKATLTADKSSSAIYMQLSLTSEDSAVYYCARRDITTFYMGOGTTLVSS 116
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 NQFKGKATLTVDTSSSPAYMQLSLTSEDSAVYYCANLRGYFDYMGOGTTLVSS 116

RESULT 6
JL0076

Ig heavy chain precursor V region (anti-phenylloxazone, 18C10) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
 C:Accession: J10076
 R:Kaartinen, M.; Rocca-Serra, J.; Maekela, O.
 M01: Immunol. 25, 859-865, 1988
 A:Title: Combinatorial association of V genes: one VH gene codes for three non-cross-
 A:Reference number: J10076; MUID:89096973; PMID:3211160

A;Residues: 1-141 <KAA>
A;Cross-references: GB:M27786; NID:g195851; PIDN:AAA38441.1; PID:g195852

F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-141/Product: Ig heavy chain #status predicted <MM

F;69-85/Region: complementarity-determining 2
F;123-135/Region: J2 segment

Query Match	78.5%;	Score 487;	DB 2;	Length 141.
Best Local Similarity	78.4%;	Pred. No. 1.2e-36;		

	Matches	91; Conservative	9; Mismatches	16; Indels	0; Gaps	0;
Q7	1	QVQLQESAEELSTKRGASVKRMSCKASGTFYFTSWLHWIRQPCQGLLEMTGYINPRADYYER	60			
Db	20	QVQLQQPAAELVYKPKPASYKLSCKRASGYFTYSWMHWIRQPCRGLEMTWGRIDNPNGSKRY	79			

Query Match	77.7%;	Score 482;	DB 1;	Length 139;
Best Local Similarity	77.5%;	Pred. No. 3.2e-36;		
Matches 93;	Conservative 10;	Mismatches 13;	Indels 4;	Gaps 2

[illegible]

RESULT 11
 C4402
 Entry leader/1g heavy chain anti-NP/1linker type 205/alkaline phosphatase fusion protein
 Species: synthetic
 Date: 06-Nov-1998 #sequence_revision 06-Nov-1998 #text_change 06-Nov-1998
 Accession: PC4402
 Suzuki, C.; Ueda, H.; Suzuki, E.; Nagamune, T.
 Biochem. 122, 332-329, 1997
 Title: Construction, bacterial expression, and characterization of haptan-specific shiga
 Reference number: PC4402
 Accession: PC4402
 Molecule type: DNA
 Residues: 1-287 <SUZ>
 Keywords: fusion protein

Query Match	Similarity	Score	DB	Length
Best Local	77.5%	Pred. No. 6,6e-36;		
Matches	93;	Conservative 10;	Mismatches 13;	Indels 4; Gaps 2
1	QVQLAESGAEISKPGASVKASCTGFTSTYWLHWIKORPGQGLEWIGYINPRNDTEY	60		
160	QVQLQDQGAELVYFGASVKLSCTASGTYFTSTYMHVHWKORPGLEWIGRIDPNSGGTKY	219		
61	NONFKRATLVADSSSTAYMQLSLTSDSAVYYCARDDI---TTF-YMGQGTTLVYSS	118		
220	NEKRSKATLVDPKPSSTAYMQQLSLTSDSAVYYICARDIYGSSYFPIWQGTTLVYSS	219		

of heavy chain V region - mouse
 Species: Mus musculus (house mouse)
 Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 Accession: S20643
 Losman, M.; Fasy, T.M.; Novick, K.E.; Monestier, M.
 Submitted to the EMBL Data Library, February 1992
 Description: Relationships among antinuclear antibodies from autoimmune MRL mice react
 Reference number: S20639
 Accession: S20643
 Status: preliminary
 Molecule type: DNA
 Residues: 1-122 <LOS>
 Cross-references: EMBL:X64998: NID:g525606: PIDW:CAA46131.1: PID:g52607
 Superfamily: Immunoglobulin V region; Immunoglobulin homology
 Keywords: heterodimer; Immunoglobulin
 5-38/Domain: Immunoglobulin homology <IMM>

Query Match	77.5%	Score	480.5	DB 2	length	122			
Best Local Similarity	74.6%	Prod. No.	3.8e-36						
Matches	91	Conservative	9	Mismatches	15	Indels	7	Gaps	1
1	QVQLQESGAEISIRPGASVYKMSKASGYTFETSWLMHVKRPGQGLEWIGYINPRNDTVEY	60							
1	QVQLQXXGAEIYKPGASVYKLSCKASGYTIRIRIHMKVRPGQGLEWIGETIDPSDNTYY	60							
61	NONFKDRTIADKSSSTAYMQLSITSDSAVYICARDITTF-----YMGCGITLL	113							
61	NONKRGATILVDKSSSTAYMQLSITSDSAVYICARREYIDLRGRAMDYMGCGISVT	120							

QY	114 VS 115
Db	121 VS 122

RESULT 13
S38717
Ig heavy chain V region - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C/Accession: S38717
R/Cimanis, A.Y.
submitted to the EMBL Data Library, November 1993
:/Reference number: S38713

A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-118 <CIN>
A: Cross references: EMBL:Y76020; NID:9416099; PIDD:CA553607.1; PIDD:9133426267
A: Superfamily: Immunoglobulin V region; Immunoglobulin homology
A: Keywords: heterotrimer; immunoglobulin
A: 15-96/Domain: Immunoglobulin homology <IM>

```

Query Match      77.1%; Score 478; DB 2; Length 118;
Best Local Similarity 77.1%; Pred. No. 6, Ie-36;
Matches    91; Conservative   12; Mismatches   13; Indels     2; Gaps     1

QY      1 OVOLSGEALSKPGASVYKMSCRKSAGTYFTSYWLHIKORPBGGLWIGINPRNDYTEY 60
        | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db       1 QPOLQGSDGLAKPKPSAVYKMSCKSAGTFTNYVIHWVKQRPGGLEICGIIHRYNDSKX 60

QY      61 NONFKDQATLADKSSSTAYMQLSLTSSEDSAVYYCARDDITTF--YWGCGTTLVVSS 116
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       61 NDRFNGKATLTLSDKSSSTAYMELSLTSEDSAVYYCARENGNFYFDWMGGTTLTVSS 118

```

RESULT 14
PH1005
Ig heavy chain V region (clone 202.54) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C/Accession: PH1005
R/Tillman, D.W.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A>Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A/Reference number: PH0971; MUID:92381444; PMID:1512540
A/Accession: PH1005
A>Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-106 <Title>
A/Experimental source: B cell, strain [NZB x NZW/F1
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
C/ID: 15-98/Domain: immunoglobulin homology <IMM>

```

Query Match: 76.8%; Score 476; DB 2; Length 106;
Best Local Similarity 89.7%; Pred. No. 8, 3e-36;
Matches 87; Conservative 6; Mismatches 4; Indels 0; Gaps 0

Cy 1 QVOLDSCGAEKSPASVYKMSCKASGTFYTSYWLHMLKOPKGGLEGWIGTYINPRNDYTEX 60
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1 QVOLDSCGAEKLPASVYKMSCKASGTFYTSYWLHMLKOPKGGLEGWIGTYINPSSDYTEX 60
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|

```

61 NQFKDKATLTADKSSSTAYMQLSLSTSEDSAYVCCA 97
 I heavy chain V region (35.8.2H) - mouse (fragment)
 C30560
 CgSpecies: Mus musculus (house mouse)
 CgDate: 23-Mar-1989 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999

RESULT 15
C30560
Ig heavy chain V region (35.8.2H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:date: 23-Mar-1989 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999

RESULT 15
C30560
Ig heavy chain V region (35.8.2H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:date: 23-Mar-1989 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999

C;Accession: C30560
R;Matsuda, T.; Kabat, E.A.

R; Matsuda, T.; Kabat, E.A.

J. Immunol. 142, 863-870, 1989

A;Title: Variable region cDNA sequences and antigen binding specificity of mouse monoclonal

A;Reference number: A30560; MUID:89110062; PMID:2464028
A;Accession: C30560

A; Accession: C30560

A;Status: preliminary
A;Molecule type: mpna

A;Residues: 1-118 <MA

A;Cross-references: GB:

C/Superfamily: Immunoglobulin V region; Immunoglobulin homology

Keywords: interleukin-15; immunoglobulin Fcγ15-98/Domain: immunoglobulin <IMM>

[illegible]

Query Match	76.68;	Score 475;
-------------	--------	------------

Query Match	76.6%	Score 475;	DB 2;	length 118;
Best Local Similarity	76.3%	Pred. No.	1.1e-35;	
Matches	90;	Conservative	12;	Mismatches 14;
				Indels 2;
				Gaps 1;

OV 1 OV0LOESGAELSKPGASVKMCKASGYTFTSYWTHWIKORPQGLFWIGYINPRNDYTEY 60

[illegible]

Db 1 QVHLQSGAELVKPGASVKISCKASGYTFTSYMNWVKQRPQGGLIEWIGEIDPSNSTNN 60

0x 61 NONEKD KATI TADKSSSTAYMOI.SITSEDSAVVYCAPBDITTE--VWGCGTTITVSS 116

[illegible]

Db 61 NQKFKNKATLTVDKSSNTAYMQLSSLTSEDSAVYYCARWGTGSWEAYWGGLTVSA 118

```
Search completed: October 7, 2003, 19:21:49
Job time : 44.2 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 23.6 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-4

Perfect score: 620
Sequence: 1 QVQLDSGAEISKPGASVKM.....ARRDTTFYGGCTLTFS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	482	77.7	139	1	HV07_MOUSE
2	465.5	75.1	120	1	HV03_MOUSE
3	456	73.5	137	1	HV11_MOUSE
4	455.5	73.5	140	1	HV02_MOUSE
5	448.5	72.3	138	1	HV48_MOUSE
6	443	71.5	120	1	HV50_MOUSE
7	435.5	70.2	117	1	HV12_MOUSE
8	434.5	70.1	117	1	HV13_MOUSE
9	432	69.7	117	1	HV04_MOUSE
10	430	69.4	117	1	HV05_MOUSE
11	429	69.2	117	1	HV06_MOUSE
12	428.5	69.1	121	1	HV01_MOUSE
13	424	68.4	117	1	HV09_MOUSE
14	421	67.9	117	1	HV49_MOUSE
15	420	67.7	117	1	HV10_MOUSE
16	415	66.9	118	1	HV51_MOUSE
17	404	65.2	117	1	HV14_MOUSE
18	392	63.2	117	1	HV15_MOUSE
19	386.5	62.3	136	1	HV15_MOUSE
20	373	60.2	147	1	HV1C_HUMAN
21	369	59.5	114	1	HV00_MOUSE
22	368	59.4	117	1	HV18_MOUSE
23	354	57.1	117	1	HV1E_HUMAN
24	329	53.1	119	1	HV17_MOUSE
25	328.5	52.9	119	1	HV41_MOUSE
26	328	52.9	119	1	HV38_MOUSE
27	320	51.6	117	1	HV40_MOUSE
28	316.5	50.9	117	1	HV42_MOUSE
29	315.5	50.7	118	1	HV39_MOUSE
30	314.5	50.4	125	1	HV1F_HUMAN
31	312.5	50.4	142	1	HV01_RAT
32	311	50.2	111	1	HV35_MOUSE
33	308	49.7	122	1	HV3G_HUMAN

34	306.5	49.4	117	1	HV17_MOUSE	P01786	mus	musculu
35	304	49.0	136	1	HV16_MOUSE	P01783	mus	musculu
36	303.5	49.0	121	1	HV37_HUMAN	P01771	homo	sapien
37	302.5	48.8	115	1	HV32_MOUSE	P01801	mus	musculu
38	299.5	48.3	113	1	HV30_MOUSE	P01799	mus	musculu
39	299	48.2	117	1	HV1A_HUMAN	P01742	homo	sapien
40	298.5	48.1	119	1	HV31_HUMAN	P01770	homo	sapien
41	298	48.1	120	1	HV1H_HUMAN	P80421	homo	sapien
42	298	48.1	124	1	HV1D_HUMAN	P01769	homo	sapien
43	297	47.9	122	1	HV3H_HUMAN	P01769	homo	sapien
44	294.5	47.5	113	1	HV27_MOUSE	P01796	mus	musculu
45	293.5	47.3	115	1	HV33_MOUSE	P01802	mus	musculu

ALIGNMENTS

```

RESULT 1
ID HV07_MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region BI-8/186-2 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Botwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THE BI-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC (NPB ANTIBODIES).
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC
CC -----
CC EMBL; J00529; AAA38170.1; -
CC PIR; A90809; MHMS18.
CC DR PDB; 1A6U; 27-MAY-98.
CC DR PDB; 1A6W; 15-JUL-98.
CC DR InterPro; IPR007110; Ig-like.
CC DR InterPro; IPR003006; Ig_MHC.
CC DR InterPro; IPR003596; Ig_V.
CC DR Pfam; PF00047; Ig; 1.
CC DR SMART; SM00406; IGV; 1.
CC DR PROSITE; PSS0835; IG_LIKE; 1.
CC KW Immunoglobulin V region; Signal; 3D-structure.
CC FT SIGNAL 1
CC FT CHAIN 1
CC FT DOMAIN 20 139 IG HEAVY CHAIN V REGION BI-8/186-2.
CC FT DOMAIN 20 49 FRAMEWORK-1.
CC FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
CC FT DOMAIN 55 68 FRAMEWORK-2.
CC FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
CC FT DOMAIN 86 117 FRAMEWORK-3.
CC FT DOMAIN 118 124 D SEGMENT.
CC FT DOMAIN 125 139 JH2 SEGMENT.
CC FT DISULFID 41 115 BY SIMILARITY.
CC FT NON_TER 139 139
CC SQ SEQUENCE 139 AA; 13419 MW; 1B57DD4FD0C9F465 CRC64;

```

```

Query Match      77.7%; Score 482; DB 1; Length 139;
Best Local Similarity 77.5%; Pred. No. 8e-42;
Matches 93; Conservative 10; Mismatches 13; Indels 4; Gaps 2;

OY 1 OVQLOESGAELEKSPGASVYKMSCKASGYTFTSYWLHMVKORPGQGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 20 OVQLOQPGAELEKVPKASVYKLSCKASGYTFTSYWMHVKORPGRLGRIIDPNSGGTKY 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 NONFKKATLTADKSSSTAYMOLSLTSEDSAVYICARDITF--TTF-YMGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 80 NEKFKSKATLTVDKSSSTAYMOLSLTSEDSAVYICARDYGGSYFDYMGQGTTLTVSS 139
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 2
HV03_MOUSE
ID HV03_MOUSE STANDARD; PRT; 120 AA.
AC P01747;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Ig heavy chain V region 36-65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RX MEDLINE=831846; PubMed=6186498;
RA Silevitz M., Gefter M.L., Brodeur P., Riblet R.,
RA Marshak-Rothstein A.;
RT "The genetic basis of antibody production: the dominant anti-arsenate
RT idotype response of the strain A mouse.";
RT Eur. J. Immunol. 12:1023-1032(1982).
CC -1- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER
CC DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE, THE AUTHORS
CC CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J
CC SEGMENT, JH2.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC HSP; P01789; IMCP.
DR InterPro: IPR007110; Ig-1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
DR Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 111
FT NON_TER 120 120
FT IG-LIKE.
SQ SEQUENCE 120 AA; 13307 MW; FFO4EAA167B654AF CRC64;

Query Match      75.1%; Score 465.5; DB 1; Length 120;
Best Local Similarity 75.0%; Pred. No. 3.1e-40;
Matches 90; Conservative 9; Mismatches 16; Indels 5; Gaps 1;

OY 2 VOLQSSGAELSPGASVYKMSCKASGYTFTSYWLHMVKORPGQGLEWIGYINPRNDYTEY 61
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 VOLQSSGAELVPGASVYKMSCKASGYTFTSYGINVVKORPGQGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 62 ONFKKATLTADKSSSTAYMOLSLTSEDSAVYICARR----DITTFYMGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 EKFKKTTLTVDKSSSTAYMOLSLTSEDSAVYICARSYVYGGSYFDYMGQGTTLTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
HV11_MOUSE
ID HV11_MOUSE STANDARD; PRT; 137 AA.
AC P01755;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE Ig heavy chain V region S43 precursor.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RX MEDLINE=81234548; PubMed=6708376;
RA Boltwell A.L.M., Paskind M., Reih M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NpB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RT Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THE GAMMA-2a CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC (NPB ANTIBODIES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; J00539; AAA38172.1; -.
CC PIR; A02038; G2MS43.
DR HSP; P01810; 2FBJ.
DR InterPro: IPR007110; Ig-1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 137
FT DOMAIN 20 49
FT DOMAIN 50 54
FT DOMAIN 55 68
FT DOMAIN 69 85
FT DOMAIN 86 117
FT DOMAIN 118 122
FT DOMAIN 123 137
FT DISULFD 41 115
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF448BEC9 CRC64;

Query Match      73.5%; Score 456; DB 1; Length 137;
Best Local Similarity 73.7%; Pred. No. 3.3e-39;
Matches 87; Conservative 12; Mismatches 17; Indels 2; Gaps 1;

OY 1 OVQLOESGAELEKSPGASVYKMSCKASGYTFTSYWLHMVKORPGQGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 20 OVQLOQPGAELEKVPKASVYKLSCKASGYTFTSYWLHMVKORPGRLGRIIDPNSGGTKY 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 NONFKKATLTADKSSSTAYMOLSLTSEDSAVYICARDITF--YMGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 80 NEKFKSKATLTVDKSSSTAYMOLSLTSEDSAVYICARILGRYDYMGQGTTLTVSS 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
HV02_MOUSE
ID HV02_MOUSE STANDARD; PRT; 140 AA.
AC P01746;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Ig heavy chain V region 93G precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RX MEDLINE=831846; PubMed=6186498;
RA Silevitz M., Gefter M.L., Brodeur P., Riblet R.,
RA Marshak-Rothstein A.;
RT "The genetic basis of antibody production: the dominant anti-arsenate
RT idotype response of the strain A mouse.";
RT Eur. J. Immunol. 12:1023-1032(1982).
CC -1- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER
CC DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE, THE AUTHORS
CC CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J
CC SEGMENT, JH2.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC HSP; P01789; IMCP.
DR InterPro: IPR007110; Ig-1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 137
FT DOMAIN 20 49
FT DOMAIN 50 54
FT DOMAIN 55 68
FT DOMAIN 69 85
FT DOMAIN 86 117
FT DOMAIN 118 122
FT DOMAIN 123 137
FT DISULFD 41 115
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF448BEC9 CRC64;

```

```

RC STRAIN-A/J;
RX MEDLINE=62152818; PubMed=6801765;
RA Sims J., Rabbits T.H., Estess P., Slaughter C., Tucker P.W.,
RA Capra J.D.;
RT "omatic mutation in genes for the variable portion of the
RT Immunoglobulin heavy chain.";
RL Science 216:309-311(1982).
CC -1- SIMILARITY: Contains 1 Immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00493; AAA38128.1; -
DR PIR: A94264; HYMSG7.
DR HSSP: P01810; ZEBU.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igy; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Hybridoma; Signal.
FT SIGNAL 1 19
FT CHAIN 20 140 IG HEAVY CHAIN V REGION 9367.
FT DOMAIN 20 139 IG-LIKE.
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15514 MW; 25A4CBBE31DA5CE8 CRC64;

Query Match 73.5%; Score 455.5; DB 1; Length 140;
Best Local Similarity 72.7%; Pred. No. 3.8e-39;
Matches 88; Conservative 9; Mismatches 19; Indels 5; Gaps 1;

QY 1 QVQLQESGAELEKPGASVKMSCKASGTFRTSYWLMHKORPGQGLEWIGYINPRNDYTEY 60
DB :|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|
20 EVQLQDSGAELEKPGASVKMSCKASGTFRTSYWLMHKORPGQGLEWIGYINPRNDYTEY 79
QY 61 NONFKDRAITLADKSSSTAAWQLSLTSESAVYVCARD--ITTFYWGQGITLVSS 115
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
80 NEFKKATITLVDKSSSTAAWQLSLTSESAVYFCARSHYGGSYDFYWGQGITLVTS 139
QY 116 S 116
DB 140 S 140

RESULT 5
HY48_MOUSE STANDARD; PRT; 138 AA.
ID HY48_MOUSE
AC P03980;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region TEPC 1017 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84248078; PubMed=6429663;
RA Gilliam A.C., Shen A., Richards J.E., Blattner F.R., Mushinski J.F.,
RA Tucker P.W.;
RT "Illegitimate recombination generates a class switch from C mu to C
RT delta in an Igd-secreting plasmacytoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4164-4168(1984).
DR PIR: A02033; HYMS77.
DR HSSP: P01810; ZEBU.
DR InterPro: IPR007110; Ig-like.

```

```

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igy; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 138 IG HEAVY CHAIN V REGION TEPC 1017.
FT DOMAIN 21 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 127 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 128 138 FRAMEWORK-4.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15576 MW; 748157E4C6907B8E CRC64;

Query Match 72.3%; Score 448.5; DB 1; Length 138;
Best Local Similarity 70.6%; Pred. No. 1.9e-38;
Matches 84; Conservative 14; Mismatches 18; Indels 3; Gaps 1;

QY 1 QVQLQESGAELEKPGASVKMSCKASGTFRTSYWLMHKORPGQGLEWIGYINPRNDYTEY 60
DB :|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|
20 QVQLQDSGAELEKPGASVKMSCKASGTFRTSYWLMHKORPGQGLEWIGYINPRNDYTEY 79
QY 61 NONFKDRAITLADKSSSTAAWQLSLTSESAVYVCARD--ITTFYWGQGITLVSS 116
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
80 NEFKKATITLVDKSSSTAAWQLSLTSESAVYFCARSDGYDFYWGQGITLVTS 138
QY 116 S 116
DB 140 S 140

RESULT 6
HY50_MOUSE STANDARD; PRT; 120 AA.
ID HY50_MOUSE
AC P06329;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region AC38 15.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=64182519; PubMed=6201362;
RA Dildrop R., Bovens J., Siekevitz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes.";
RL EMBO J. 3:517-523(1984).
DR PIR: A02037; MHMS15.
DR HSSP: P01810; ZEBU.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igy; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT SIGNAL 1 98
FT CHAIN 99 105 V SEGMENT.
FT DOMAIN 106 120 D SEGMENT.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 914453F426F09834 CRC64;

Query Match 71.5%; Score 443; DB 1; Length 120;
Best Local Similarity 71.7%; Pred. No. 5.8e-38;
Matches 86; Conservative 9; Mismatches 21; Indels 4; Gaps 2;

QY 1 QVQLQESGAELEKPGASVKMSCKASGTFRTSYWLMHKORPGQGLEWIGYINPRNDYTEY 60
DB :|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|

```

```

Db      1 OVQLOESGAEISKPGASVNLSCKASGYTFTSYMHMIRQPGGLEWIGINPNSNGTNY 60
Qy      61 NONFKDKATLTADKSSSTAYVMQSLTSDSAVYVCARDITF--TTFY--WGOGTTLTVSS 116
      1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 NEKFKKATLTVDKSSSTAYVMQSLTSDSAVYVCARDYEGDRYFDVWGAGTTVTYSS 120

RESULT 7
HVL2_MOUSE
ID      HVL2_MOUSE      STANDARD:      PRT;      117 AA.
AC      P01756;
DR      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DE      15-SEP-2003 (Rel. 42, Last annotation update)
OS      Ig heavy chain V region MOPC 104E.
OC      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      [1]
      NCBI_TaxID=10090;
      SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE ASN-55.
RX      MEDLINE=83075344; PubMed=6816276;
RA      Kehy M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
      Hood L.E.;
RT      "Complete amino acid sequence of a mouse mu chain: homology among
      heavy chain constant region domains."
      Biochemistry 21:5415-5424(1982).
CC      -1- MISCELLANEOUS: THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA
      PROTEIN HAS ALSO BEEN DETERMINED.
CC      -1- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR      PIR: A02039; MMSAE.
DR      HSSP: P01789; 1MCP.
DR      InterPro: IPR007110; Ig-like.
DR      InterPro: IPR003006; Ig_MHC.
DR      InterPro: IPR003596; Ig_V.
DR      Pfam: PF00047; Ig_1.
DR      SMART: SM00406; IGV: 1.
DR      PROSITE: PS50835; IG_LIKE; 1.
KW      Immunoglobulin V region; Glycoprotein.
FT      DOMAIN 1 116
FT      DISULFID 22 96
FT      CARBOHYD BY SIMILARITY.
FT      NON_TER 117 117
FT      N-LINKED (GLCNAC. . .) (COMPLEX).
SQ      SEQUENCE 117 AA; 12983 MW; 3CF8ACE4BE47E41 CRC64;

Query Match      70.2%; Score 435.5; DB 1; Length 117;
Best Local Similarity 72.9%; Pred. No. 3, 2e-37;
Matches 86; Conservative 10; Mismatches 19; Indels 3; Gaps 2;

Qy      1 OVQLOESGAEISKPGASVNLSCKASGYTFTSYMHMIRQPGGLEWIGINPNSNGTNY 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 EVQLOOSGELVLPKASVYKMSCKASGYTFTDYMKVWVSKSLKLEWIGIDINPNNGGTSY 60

Qy      61 NONFKDKATLTADKSSSTAYVMQSLTSDSAVYVCARDITTFY--WGOGTTLTVSS 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 NOKFKKATLTVDKSSSTAYVMQSLTSDSAVYVICA-RDYMYEDVWGAGTTVTYSS 117

RESULT 8
HVL3_MOUSE
ID      HVL3_MOUSE      STANDARD:      PRT;      117 AA.
AC      P01757;
DR      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DE      15-SEP-2003 (Rel. 42, Last annotation update)
OS      Ig heavy chain V region J558.
OC      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
      [1]
      SEQUENCE.

```

```

RX      MEDLINE=80078170; PubMed=6765983;
RA      Schilling J., Clevinger B., Davie J.M., Hood L.;
RT      "Amino acid sequence of homogeneous antibodies to dextran and DNA
      rearrangements in heavy chain V-region gene segments."
      Nature 283:35-40(1980).
CC      -1- MISCELLANEOUS: THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO
      BIND DEXTRAN DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF
      WHICH OCCUR IN THE D AND J SEGMENTS.
CC      -1- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR      PIR: A26242; MMSJ5.
DR      HSSP: P01789; 1MCP.
DR      InterPro: IPR007110; Ig-like.
DR      InterPro: IPR003006; Ig_MHC.
DR      InterPro: IPR003596; Ig_V.
DR      Pfam: PF00047; Ig_1.
DR      SMART: SM00406; IGV: 1.
DR      PROSITE: PS50835; IG_LIKE; 1.
KW      Immunoglobulin V region.
FT      DOMAIN 1 116
FT      DISULFID 22 96
FT      NON_TER 117 117
FT      BY SIMILARITY.
SQ      SEQUENCE 117 AA; 13024 MW; 292E2AF4BE47E41 CRC64;

Query Match      70.1%; Score 434.5; DB 1; Length 117;
Best Local Similarity 73.5%; Pred. No. 4, 1e-37;
Matches 86; Conservative 8; Mismatches 22; Indels 1; Gaps 1;

Qy      1 OVQLOESGAEISKPGASVNLSCKASGYTFTSYMHMIRQPGGLEWIGINPNSNGTNY 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 EVQLOOSGELVLPKASVYKMSCKASGYTFTDYMKVWVSKSLKLEWIGIDINPNNGGTSY 60

Qy      61 NONFKDKATLTADKSSSTAYVMQSLTSDSAVYVCARDITTFY--WGOGTTLTVSS 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 NOKFKKATLTVDKSSSTAYVMQSLTSDSAVYVCARDRYMYFDVWGAGTTVTYSS 117

RESULT 9
HVL4_MOUSE
ID      HVL4_MOUSE      STANDARD:      PRT;      117 AA.
AC      P01748;
DR      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DE      15-JUL-1999 (Rel. 38, Last annotation update)
OS      Ig heavy chain V region 23 precursor.
OC      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
      [1]
      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6;
RA      MEDLINE=81234548; PubMed=6788376;
RA      Baltimore D.;
      Baltimore A.L.M., Paskind M., Rath M., Imanishi-Karl T., Rajewsky K.,
      "Heavy chain variable region contribution to the NpB family of
      antibodies: somatic mutation evident in a gamma 2a variable region."
      Cell 24:625-637(1981).
RT      Cell 24:625-637(1981).
CC      -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
      RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
DR      PIR: A02030; HVM523.
DR      HSSP: P01810; 2EBJ.
DR      InterPro: IPR007110; Ig-like.
DR      InterPro: IPR003006; Ig_MHC.
DR      InterPro: IPR003596; Ig_V.
DR      Pfam: PF00047; Ig_1.
DR      SMART: SM00406; IGV: 1.
DR      PROSITE: PS50835; IG_LIKE; 1.
KW      Immunoglobulin V region; Signal.
FT      SIGNAL 1 19
FT      CHAIN 20 117
FT      DOMAIN 20 49
FT      DOMAIN 50 54
      IG HEAVY CHAIN V REGION 23.
      FRAMEWORK-1.
      COMPLEMENTARITY-DETERMINING-1.

```

FT DOMAIN 55 68 FRAMEWORK-2.
 FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
 FT DISULFID 86 117 FRAMEWORK-3.
 FT NON_TER 41 115 BY SIMILARITY.
 FT NON_TER 117 117
 SO SEQUENCE 117 AA; 12772 MW; C530F82C906F69B CRC64;
 Query Match 69.7%; Score 432; DB 1; Length 117;
 Best Local Similarity 81.6%; Pred. No. 7.3e-37;
 Matches 80; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
 QY 1 QVLOESGAELEKPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGINPRNDYTEY 60
 DB 20 QVLOQPGAELEVRPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGINPRNDYTEY 79
 QY 61 NONFKDKATLTADKSSSTAYMQLSLSLSEDSAVYYCAR 98
 DB 80 NEKFKSVTLTVDKSSSTAYTQSLSLSEDSAVYYCAR 117
 RESULT 10
 HV05_MOUSE STANDARD; PRT; 117 AA.
 AC P01749;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ig heavy chain V region 3 precursor.
 GN IGH-VJ558.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=81234548; PubMed=6788376;
 RA Botwell A.L.M., Paskind M., Reith M., Imanishi-Kari T., Rajewsky K.,
 Baltimore D.;
 RT "Heavy chain variable region contribution to the NpB family of
 RT antibodies: somatic mutation evident in a gamma 2a variable region.";
 RL Cell 24:625-637(1981).
 CC -I- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
 CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J00536; AAA38605.1; -
 DR PIR: A02031; HVMS3.
 DR HSSP: P01810; 2PBJ.
 DR MGD; MGI:96466; Igh-VJ558.
 DR InterPro: IPR007110; Igh-V-like.
 DR InterPro: IPR003006; Igh-MHC.
 DR InterPro: IPR003596; Igh-V.
 DR Pfam: PF00047; Igh; 1.
 DR SMART: SM00406; Igh; 1.
 DR PROSITE: PS50835; Igh-Like; 1.
 DR Immunoglobulin V region; signal.
 FT CHAIN 1 19
 FT SIGNAL 19
 FT DOMAIN 20 49 IG HEAVY CHAIN V REGION 3.
 FT DOMAIN 50 54 FRAMEWORK-1.
 FT DOMAIN 55 58 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 69 85 FRAMEWORK-2.
 FT DOMAIN 86 115 COMPLEMENTARITY-DETERMINING-2.
 FT DISULFID 41 115 FRAMEWORK-3.
 FT NON_TER 117 117 BY SIMILARITY.

SO SEQUENCE 117 AA; 13016 MW; 427C861C53975EDC CRC64;
 Query Match 69.4%; Score 430; DB 1; Length 117;
 Best Local Similarity 81.6%; Pred. No. 1.2e-36;
 Matches 80; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
 QY 1 QVLOESGAELEKPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGINPRNDYTEY 60
 DB 20 QVLOQPGAELEVRPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGINPRNDYTEY 79
 QY 61 NONFKDKATLTADKSSSTAYMQLSLSLSEDSAVYYCAR 98
 DB 80 NEKFKSVTLTVDKSSSTAYTQSLSLSEDSAVYYCAR 117
 RESULT 11
 HV06_MOUSE STANDARD; PRT; 117 AA.
 AC P01750;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig heavy chain V region 102 precursor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=81234548; PubMed=6788376;
 RA Botwell A.L.M., Paskind M., Reith M., Imanishi-Kari T., Rajewsky K.,
 Baltimore D.;
 RT "Heavy chain variable region contribution to the NpB family of
 RT antibodies: somatic mutation evident in a gamma 2a variable region.";
 RL Cell 24:625-637(1981).
 CC -I- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
 CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J00536; AAA38605.1; -
 DR PIR: A02032; HVMS2.
 DR HSSP: P01810; 2PBJ.
 DR InterPro: IPR007110; Igh-V-like.
 DR InterPro: IPR003006; Igh-MHC.
 DR InterPro: IPR003596; Igh-V.
 DR Pfam: PF00047; Igh; 1.
 DR SMART: SM00406; Igh; 1.
 DR PROSITE: PS50835; Igh-Like; 1.
 DR Immunoglobulin V region; signal.
 FT CHAIN 1 19
 FT SIGNAL 19
 FT DOMAIN 20 49 IG HEAVY CHAIN V REGION 102.
 FT DOMAIN 50 54 FRAMEWORK-1.
 FT DOMAIN 55 58 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 69 85 FRAMEWORK-2.
 FT DOMAIN 86 115 COMPLEMENTARITY-DETERMINING-2.
 FT DISULFID 41 115 FRAMEWORK-3.
 FT NON_TER 117 117 BY SIMILARITY.
 SO SEQUENCE 117 AA; 12867 MW; 740A65DB851CABC CRC64;
 Query Match 69.2%; Score 429; DB 1; Length 117;
 Best Local Similarity 83.3%; Pred. No. 1.5e-36;
 Matches 80; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 QY 2 VLOESGAELEKPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGINPRNDYTEY 61
 DB 21 VLOQPGAELEVRPGASVYKSCASGYFTSYMLHWIKORPGGLEWIGIRHPSDNYN 80
 QY 62 QNFADKATLTADKSSSTAYMQLSLSLSEDSAVYYCA 97
 DB 81 QNFKATLTADKSSSTAYMQLSLSLSEDSAVYYCA 116
 RESULT 12
 HV01_MOUSE

ID	HPV1_MOUSE	STANDARD:	PRT:	121 AA.
AC	P01745;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig heavy chain V region MPC11.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=81053741; PubMed=6253904;			
RA	zakut R., Cohen J., Givol D.;			
RT	"Cloning and sequence of the cDNA corresponding to the variable			
RL	region of Immunoglobulin heavy chain MPC11."			
RL	Nucleic Acids Res. 8:3591-3601(1980).			
RN	[2]			
RP	REVISIONS.			
RA	zakut R., Cohen J., Givol D.;			
RL	Nucleic Acids Res. 8:4839-4840(1980).			
CC	-1- MISCELLANEOUS: THIS SEQUENCE WAS TRANSLATED FROM AN MRNA ISOLATED			
CC	FROM A MELOMA THAT SECRETES IGG2B.			
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.			
CC	PIR: A93708; PGM81.			
DR	HSSP: P01810; 2P8J.			
DR	InterPro: IPR007110; Ig-like.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; Ig; 1.			
DR	SMART: SM00406; Igv; 1.			
DR	PROSITE: PS50835; IG-LIKE; 1.			
KW	Immunoglobulin V region.			
FT	DOMAIN	1	112	IG-LIKE.
FT	NON_TER	121	121	
SO	SEQUENCE	121 AA;	13135 MM;	227AEF3EC56ED0BF CRC64;
Query Match		69.1%;	Score 428.5;	DB 1: Length 121;
Best Local Similarity		66.1%;	Pred. No. 1.7e-36;	
Matches	82;	Conservative 14;	Mismatches 17;	Indels 11; Gaps 2
QY	1 QVQLQESGAEELSKPGASVYKMSCKASGYFTSYWLMHMQRPQGQGLWIGIYINPRNDYFEY			60
DB	1 EAQLQGSAGELVRRGTSTKVISCKAKAGYFTNWMVGWVERGHGLEMIGDIYPGGFTNY			60
QY	61 NQNFRRKATLTADSSSTAVYMQLSLSIEDSAVYVCARRDITFEY-----WGQFTL			112
DB	61 NDNKTKGKTLTADPTSSSTAVYQLSLTSEDSALYHCAR---GIYNSSPYFDMSWGQFTL			117
QY	113 TVSS 116			
DB	118 TVSS 121			
RESULT 13				
HPV09_MOUSE				
ID	HPV09_MOUSE	STANDARD:	PRT:	117 AA.
AC	P01753; P11271;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig heavy chain V region 18c-1 precursor.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RX	MEDLINE=81234548; PubMed=6788376;			
RA	Bothwell A.L.M., Paekind M., Reth M., Imanishi-Kari T., Rajewsky K.,			
RA	Baltimore D.;			
RT	"Heavy chain variable region contribution to the Npb family of			

```

antibodies: somatic mutation evident in a gamma 2a variable region."
Cell 24:625-637(1981).
-1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
CC PIR: D90809; HVMS61.
DR HSSP; P01810; 2PBJ.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR PIRam; PF00047; Ig; 1.
DR SMART; SMO0406; IgV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 186-1.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
SQ SEQUENCE 117 AA; 12890 MM; 16191A088CB17F5A CRC64;

Query Match 68.4%; Score 424; DB 1; Length 117;
Best Local Similarity 79.6%; Pred. No. 4.7e-36;
Matches 78; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVQLQESGAELTKSKGASVSKMSCKASGYTFTSYMLHWIKORPGOGLEWIGYINPRNDYTEY 60
Db 20 QVQLQDQGAELVYKKGASVSKLSCKASGYTFTSYMMHWIKORPGOLEWIGRIDRNSGSKTY 79
QY 61 NQNFKDKATLTADKSSSTAYWQSSLTSEDSAVYYCAR 98
Db 80 NEKFKSKATLTVDTSSTAYWQHLSTSEDSAVYYCAR 117

RESULT 14
ID HV49_MOUSE STANDARD; PRT; 117 AA.
AC P06328;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH58 B4 precursor.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D.; Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
unrearranged VH gene segments.";
RL Cell 40:271-281(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M13788; AAA36506.1; -.
DR PIR; A02035; MHMSB4.
DR HSSP; P01810; 2PBJ.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SMO0406; IgV; 1.

```


DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION VH558 B4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12834 MW; B8862FAC67ABD345 CRC64;
Query Match 67.9%; Score 421; DB 1; Length 117;
Best Local Similarity 79.4%; Pred. No. 9.4e-36;
Matches 77; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 2 VOLESGBALSLKPGASVYKASGTYFTSYLHWIKORPGGLEWIGTINPRNDYTEY 61
DB 21 VOLESGBALSLKPGASVYKASGTYFTSYLHWIKORPGGLEWIGTINPRNDYTEY 80
QY 62 ONEFKKATLTADKSSSTAYWOLSLTSEDSAVYYCAR 98
DB 81 EKFESKATLTVDKPSSTAYWOLSLTSEDSAVYYCTR 117
RESULT 15
ID HV10_MOUSE STANDARD; PRT; 117 AA.
AC P01754; P11270;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE I9 heavy chain V region 145 precursor.
GN IGH-VJ558.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Botlwell A.V.M., Paskind M., Reith M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERM LINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00533; AAA38602.1; -.
DR PIR: C90809; HVMS45.
DR HSSP: P01810; 2FBU.
DR MGD: MGI:96486; Igh-VJ558.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 145.
FT DOMAIN 20 49 FRAMEWORK-1.

FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12921 MW; D37DE8A3F543E996 CRC64;
Query Match 67.7%; Score 420; DB 1; Length 117;
Best Local Similarity 79.6%; Pred. No. 1.2e-35;
Matches 78; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY 1 VOLESGBALSLKPGASVYKASGTYFTSYLHWIKORPGGLEWIGTINPRNDYTEY 60
DB 20 VOLESGBALSLKPGASVYKASGTYFTSYLHWIKORPGGLEWIGTINPRNDYTEY 79
QY 61 ONEFKKATLTADKSSSTAYWOLSLTSEDSAVYYCAR 98
DB 80 NEFKSKATLTVDKPSSTAYWOLSLTSEDSAVYYCTR 117
Search completed: October 7, 2003, 19:15:00
Job time : 24.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 115.6 Seconds
(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-4

Perfect score: 620
Sequence: 1 OVQLOESGALSRPGASVKRM.....ARRDITFYWGQITLTVSS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MNC:*
8: SP_ORNITHINE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	491.5	79.3	142	11	0924Q1 mus musculus
2	490	79.0	141	11	0924Q4 mus musculus
3	489	78.9	145	11	0924Q9 mus musculus
4	487	78.5	137	11	0924R6 mus musculus
5	485	78.1	145	11	0924R1 mus musculus
6	484.5	78.1	146	11	0924Q8 mus musculus
7	483	77.9	145	11	0924P7 mus musculus
8	482	77.7	145	11	0924Q7 mus musculus
9	480	77.4	143	11	0924P9 mus musculus
10	479	77.3	143	11	0924R0 mus musculus
11	477.5	77.0	140	11	0924P8 mus musculus
12	477	76.9	145	11	0924R4 mus musculus
13	476.5	76.9	146	11	0924R8 mus musculus
14	476.5	76.9	146	11	0924Q3 mus musculus
15	473.5	76.4	142	11	0924Q2 mus musculus

17	469.5	75.7	140	11	0924R2 mus musculus
18	468	75.5	482	11	08K172 mus musculus
19	467	75.3	143	11	0924Q5 mus musculus
20	467	75.3	143	11	091V67 mus musculus
21	466	75.2	143	11	0924R7 mus musculus
22	464.5	74.9	144	11	0924P5 mus musculus
23	464	74.8	143	11	0924Q0 mus musculus
24	462	74.5	118	11	0921C4 mus musculus
25	462	74.5	139	11	0924R5 mus musculus
26	462	74.5	145	11	0924R3 mus musculus
27	461	74.4	143	11	091VA2 mus musculus
28	459	74.0	278	11	0921K1 mus musculus
29	458	73.9	109	11	09JL75 mus musculus
30	457	73.7	143	11	0924P6 mus musculus
31	454.5	73.3	117	11	09QXFO mus musculus
32	453.5	73.1	481	11	091WT3 mus musculus
33	452.5	73.0	473	11	09DBL4 mus musculus
34	452	72.9	468	11	08KOF2 mus musculus
35	450.5	72.7	481	11	091WT1 mus musculus
36	448	72.3	613	11	08VCX7 mus musculus
37	444.5	71.7	117	11	09QXEX mus musculus
38	443	71.5	473	11	099L25 mus musculus
39	436.5	70.4	119	5	09GYZ2 schistosoma
40	436	70.3	114	11	09JLB1 mus musculus
41	435.5	70.2	489	11	08VCX4 mus musculus
42	432.5	69.8	117	11	0921C6 mus musculus
43	431.5	69.6	168	11	08VDC9 mus musculus
44	429.5	69.3	480	11	08K0Z4 mus musculus
45	425	68.5	463	11	099LC4 mus musculus

ALIGNMENTS

RESULT 1

ID 0924Q1 PRELIMINARY; PRT; 142 AA.
AC 0924Q1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE V23-D-J-C mu protein (fragment).
GN V23-D-J-C MO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP)."
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069913; BAB63929.1; -
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 142
SQ SEQUENCE 142 AA; 15622 MW; 24A265CE4EA4318B CRC64;

Query Match 79.3%; Score 491.5; DB 11; Length 142;
Best local similarity 79.5%; Pred. No. 1.9e-44;

Matches 93; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 OVQLOESGALSRPGASVKRSGTFTSYLHWIKRPGGGLIEWIGTINRNDYTEY 60
DB 1 OVQLOQPGTELVRPGASVKLSKRASGTYFTSYMHWIKRPGGGLIEWIGTINRNDYTEY 60

QY 61 NONFKDKATLTPADKSSSTAYMQLSLTSEDSAVYYCARDITTF-YMGOGTTLTVSS 116
 DB 61 NEKFSKATLTVDKSSSTAYMQLSLTSEDSAVYYCARDIGWMDYMGOGTSVTVSS 117

RESULT 2

Q92404 PRELIMINARY: PRT: 141 AA.
 AC 092404: 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE VH186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB067796; BAB63281.1; -
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1 141
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA: 15561 MW: DDD08482D66876A0 CRC64;

Query Match 79.0%; Score 490; DB 11; Length 141;
 Best Local Similarity 78.4%; Pred. No. 2.7e-44;
 Matches 91; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 QVQLDSGAEISKSPASVYKMSCKASGYTFSTYWLHMIRKORPGGLEWIGYINPRNDYTEY 60
 DB 1 QVQLDQPGAEELVKKPASVYKLSCKASGYTFSTYWMHVMVQRGRGLEWIGRIDPNSGGTKY 60
 QY 61 NONFKDKATLTPADKSSSTAYMQLSLTSEDSAVYYCARDITTFYMGOGTTLTVSS 116
 DB 61 NEKFSKATLTVDKSSSTAYMQLSLTSEDSAVYYCARDYGRTPMGOGTTLTVSS 116

RESULT 3

Q92409 PRELIMINARY: PRT: 145 AA.
 AC 092409: 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE VH186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB067791; BAB63276.1; -
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003006; IG_MHC.

DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; IGV_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1 145
 FT NON_TER 145 145
 SQ SEQUENCE 145 AA: 16001 MW: 0F409EB09FA333D2 CRC64;

Query Match 78.9%; Score 489; DB 11; Length 145;
 Best Local Similarity 76.7%; Pred. No. 3.6e-44;
 Matches 92; Conservative 12; Mismatches 12; Indels 4; Gaps 1;

QY 1 QVQLDSGAEISKSPASVYKMSCKASGYTFSTYWLHMIRKORPGGLEWIGYINPRNDYTEY 60
 DB 1 QVQLDQPGAEELVKKPASVYKLSCKASGYTFSTYWMHVMVQRGRGLEWIGRIDPNSGGTKY 60
 QY 61 NONFKDKATLTPADKSSSTAYMQLSLTSEDSAVYYCARDITTF---YMGOGTTLTVSS 116
 DB 61 NEKFSKATLTVDKSSSTAYMQLSLTSEDSAVYYCARSLLTTYAMDYMGOGTSVTVSS 120

RESULT 4

Q924R6 PRELIMINARY: PRT: 137 AA.
 AC 0924R6: 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE VH186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB067783; BAB63268.1; -
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; IGV_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1 137
 FT NON_TER 137 137
 SQ SEQUENCE 137 AA: 5C38D966DC6A4124 CRC64;

Query Match 78.5%; Score 487; DB 11; Length 137;
 Best Local Similarity 79.3%; Pred. No. 5.5e-44;
 Matches 92; Conservative 9; Mismatches 11; Indels 4; Gaps 1;

QY 1 QVQLDSGAEISKSPASVYKMSCKASGYTFSTYWLHMIRKORPGGLEWIGYINPRNDYTEY 60
 DB 1 QVQLDQPGAEELVKKPASVYKLSCKASGYTFSTYWMHVMVQRGRGLEWIGRIDPNSGGTKY 60
 QY 61 NONFKDKATLTPADKSSSTAYMQLSLTSEDSAVYYCARDITTFYMGOGTTLTVSS 116
 DB 61 NEKFSKATLTVDKSSSTAYMQLSLTSEDSAVYYCARDM---YMGOGTTLTVSS 112

RESULT 5

Q924R1 PRELIMINARY: PRT: 145 AA.
 AC 0924R1: 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE VH186.2-D-J-C mu protein (Fragment).

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP)."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB067789; BAB63274.1; -
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; Igy_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 15979 MW; 0162D0A26C746C04 CRC64;

Query Match 78.2%; Score 485; DB 11; Length 145;
 Best Local Similarity 75.8%; Pred. No. 9.6e-44;
 Matches 91; Conservative 13; Mismatches 12; Indels 4; Gaps 1;
 QY 1 QVOLOESGAEILSRPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGYINPNDTXY 60
 Db 1 QVOLOPGAEILVPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGRIDPNSGGTKY 60
 QY 61 NONFKRATLTADKSSSTAYWQLSLTSEDSAVYYCARRDITTF---YWGQGTTLTVSS 116
 Db 61 NEKFKSKATLTVDKPSSTAYWQLSLTSEDSAVYYCARYDGGSPYADYWGQGTSTVYSS 120

RESULT 6
 Q92408 PRELIMINARY; PRT; 146 AA.
 AC Q92408;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE VHL186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP)."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB067792; BAB63277.1; -
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; Igy_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1
 FT NON_TER 146
 SQ SEQUENCE 146 AA; 16023 MW; 4B04959991D49159 CRC64;

Query Match 78.1%; Score 484.5; DB 11; Length 146;
 Best Local Similarity 76.9%; Pred. No. 1.1e-43;
 Matches 93; Conservative 9; Mismatches 14; Indels 5; Gaps 1;
 QY 1 QVOLOESGAEILSRPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGYINPNDTXY 60
 Db 1 QVOLOPGAEILVPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGRIDPNSGGTKY 60

Db 1 QVOLOPGAEILVPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGRIDPNSGGTKY 60
 QY 61 NONFKRATLTADKSSSTAYWQLSLTSEDSAVYYCARRDITTF-----YWGQGTTLTVSS 115
 Db 61 NEKFKSKATLTVDKPSSTAYWQLSLTSEDSAVYYCARTFTTYVAEPDYGQGTTLTVSS 120
 QY 116 S 116
 Db 121 S 121

RESULT 7
 Q92407 PRELIMINARY; PRT; 145 AA.
 AC Q92407;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE VHL186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 Affinity Maturation of B Cell Antigen Receptors in Response to (4-
 RT Hydroxy-3-Nitrophenyl)Acetyl (NP)."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB069918; BAB63934.1; -
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; Igy_1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 15988 MW; FB73958704796C9A CRC64;

Query Match 77.9%; Score 483; DB 11; Length 145;
 Best Local Similarity 74.2%; Pred. No. 1.6e-43;
 Matches 89; Conservative 16; Mismatches 11; Indels 4; Gaps 1;
 QY 1 QVOLOESGAEILSRPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGYINPNDTXY 60
 Db 1 QVOLOPGAEILVPGASVYKSKASGYTFTSYNLMHKORPGOLEMIGRIDPNSGGTKY 60
 QY 61 NONFKRATLTADKSSSTAYWQLSLTSEDSAVYYCARRDITTF---YWGQGTTLTVSS 116
 Db 61 NEKFKSKATLTVDKPSSTAYWQLSLTSEDSAVYYCARKDSSPYADYWGQGTSTVYSS 120

RESULT 8
 Q92407 PRELIMINARY; PRT; 145 AA.
 AC Q92407;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE VHL186.2-D-J-C mu protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Kozono Y., Kozono H., Azuma T.;
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
 Affinity Maturation of B Cell Antigen Receptors in Response to (4-

```
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
DR EMBL; AB067793; BAB63278.1; -.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16141 MW; 55A59A7908B2CD6A CRC64;

Query Match
Best Local Similarity 77.7%; Score 482; DB 11; Length 145;
Matches 93; Conservative 10; Mismatches 13; Indels 4; Gaps 2;

OY 1 OVQLOESGAELSKPGASVYKSCASGYFTSYLWIMIKRPGGLEWIGYINPRNDYTY 60
DB 1 OVQLOQPGALVKPGASVLSCKASGYFTSYLWIMHVKRPGRGLEWIRIDPNSGITY 60
OY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYVCARRDI--TTF-YWGQGTTLTVSS 116
DB 61 NEKFKSKATLTVDKPSSTAYMQLSLTSDSAVYVCARDYDYGSSYFDYWGQGTTLTVSS 120

RESULT 9
OY24P9 PRELIMINARY; PRT; 143 AA.
AC O924P9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE V303-D-D-C mu protein (Fragment).
GN V303-D-D-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069916; BAB63932.1; -.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15704 MW; 43CD8C72D52134F6 CRC64;

Query Match
Best Local Similarity 77.4%; Score 480; DB 11; Length 143;
Matches 92; Conservative 8; Mismatches 16; Indels 2; Gaps 1;

OY 1 OVQLOESGAELSKPGASVYKSCASGYFTSYLWIMIKRPGGLEWIGYINPRNDYTY 60
DB 1 OVQLOQPGALVKPGASVLSCKASGYFTSYLWIMHVKRPGRGLEWIRIDPNSYTY 60
OY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYVCARR--DITFFYWGQGTTLTVSS 116
DB 61 NOKFKGKATLTVDKSSSTAYMQLSLTSDSAVYVCASHYSSSDYWGQGTTLTVSS 118

RESULT 10
OY24R0
```

```
ID O924R0 PRELIMINARY; PRT; 143 AA.
AC O924R0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VHI86.2-D-D-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067790; BAB63275.1; -.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15868 MW; 139B2E96B81E07F CRC64;

Query Match
Best Local Similarity 76.3%; Score 479; DB 11; Length 143;
Matches 90; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

OY 1 OVQLOESGAELSKPGASVYKSCASGYFTSYLWIMIKRPGGLEWIGYINPRNDYTY 60
DB 1 OVQLOQPGALVKPGASVLSCKASGYFTSYLWIMHVKRPGRGLEWIRIDPNSGITY 60
OY 61 NONFKDKATLTADKSSSTAYMQLSLTSDSAVYVCARR--RDITFFYWGQGTTLTVSS 116
DB 61 NEKFKSKATLTVDKPSSTAYMQLSLTSDSAVYVCARDYDYMWDYWGQGTTLTVSS 118

RESULT 11
OY24O6 PRELIMINARY; PRT; 145 AA.
AC O924O6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VHI86.2-D-D-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067794; BAB63279.1; -.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16011 MW; 9BC0846D40DF97EA CRC64;
```

```

Query Match          77.3%; Score 479; DB 11; Length 145;
Best Local Similarity 74.2%; Pred. No. 4.2e-43;
Matches 89; Conservative 14; Mismatches 13; Indels 4; Gaps 1;

QY 1 0VOLOESGAEISKPGASVKMSCKASGYFTSYWLMHKORPGGLEWIGYINPRNDYEX 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 0VOLOQPGAEILVKGASVYKLSCKASGYFTSYWMMHWKORPGGLEWIGRIDPNSGGTKY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NONFKDKATLTPADKSSSTAYVQSLSTSEDSAVYYCAR---YWGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 NEKFKSKATLTVDKSSSTAYVQSLSTSEDSAVYYCARSLTSHYADYMGQGTSTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
Q924P8 PRELIMINARY; PRT; 140 AA.
ID 0924P8
AC 0924P8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE V23-D-J-C mu protein (Fragment).
GN V23-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA "Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB069917; BAB63933.1; -
DR InterPro; IPR007110; Ig_1Ike.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 140 AA; 15392 MW; 904C80C82548C936 CRC64;

Query Match          77.0%; Score 477.5; DB 11; Length 140;
Best Local Similarity 79.5%; Pred. No. 5.7e-43;
Matches 93; Conservative 5; Mismatches 16; Indels 3; Gaps 2;

QY 1 0VOLOESGAEISKPGASVKMSCKASGYFTSYWLMHKORPGGLEWIGYINPRNDYEX 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 0VOLOQPGAEILVKGASVYKLSCKASGYFTSYWMMHWKORPGGLEWIGINPNSGCTNY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NONFKDKATLTPADKSSSTAYVQSLSTSEDSAVYYCAR-DITTFYMGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 NEKFKSKATLTVDKSSSTAYVQSLSTSEDSAVYYCARNDPFD--YWGQGTTLTVSS 115
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
Q924R4 PRELIMINARY; PRT; 145 AA.
ID 0924R4
AC 0924R4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VHI86.2-D-J-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;

```

```

RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB067785; BAB63270.1; -
DR InterPro; IPR007110; Ig_1Ike.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 145 AA; 16081 MW; ECD81A135E05B8AA CRC64;

Query Match          76.9%; Score 477; DB 11; Length 145;
Best Local Similarity 75.0%; Pred. No. 6.8e-43;
Matches 90; Conservative 11; Mismatches 15; Indels 4; Gaps 1;

QY 1 0VOLOESGAEISKPGASVKMSCKASGYFTSYWLMHKORPGGLEWIGYINPRNDYEX 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 0VOLOQPGAEILVKGASVYKLSCKASGYFTSYWMMHWKORPGGLEWIGRIDPNSGGTKY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NONFKDKATLTPADKSSSTAYVQSLSTSEDSAVYYCAR---DITTFYMGQGTTLTVSS 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 NEKFKSKATLTVDKSSSTAYVQSLSTSEDSAVYYCARSDYDYADYMGQGTSTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
Q924R8 PRELIMINARY; PRT; 146 AA.
ID 0924R8
AC 0924R8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VHI86.2-D-J-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB067781; BAB63266.1; -
DR InterPro; IPR007110; Ig_1Ike.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 146 AA; 16216 MW; 92460F1FDF1B7538 CRC64;

Query Match          76.9%; Score 476.5; DB 11; Length 146;
Best Local Similarity 76.0%; Pred. No. 7.7e-43;
Matches 92; Conservative 10; Mismatches 14; Indels 5; Gaps 2;

QY 1 0VOLOESGAEISKPGASVKMSCKASGYFTSYWLMHKORPGGLEWIGYINPRNDYEX 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 0VOLOQPGAEILVKGASVYKLSCKASGYFTSYWMMHWKORPGGLEWIGRIDPNSGGTKY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NONFKDKATLTPADKSSSTAYVQSLSTSEDSAVYYCAR---RDITTFYMGQGTTLTVSS 115
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 NEKFKSKATLTVDKSSSTAYVQSLSTSEDSAVYYCARSYSSLYDYMGQGTTLTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 S 116

```

Db 121 S 121

RESULT 15

```
092403 PRELIMINARY; PRT; 146 AA.
AC 092403;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE VHL86.2-D-D-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Mitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB067797; BAB63282.1;
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 146
SQ SEQUENCE 146 AA; 16136 MW; CEA8D6E1955807F CRC64;
```

Query Match

76.9%; Score 476.5; DB 11; Length 146;

Best Local Similarity 74.4%; Pred. No. 7, 7e-43; Matches 90; Conservative 11; Mismatches 15; Indels 5; Gaps 1;

```
QY 1 QVQLQESGAEISKPGASVMSCKASGYFTSYWLHWIKORPGGLEWIGYINPRNDYTEY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLQPGAEIYKPGASVKLSCKASGYFTSYWMHWVKQPGRGLEWIGRIDPNSGCTKY 60
QY 61 NQNFKDKATLTADKSSSTAYMQLSSITSDSAVYICARR-----DITTFYWGQITLTVS 115
    |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NEKFSKATLTVDKPSSTAYMQLSSITSDSAVYICARSLYDYGADYAMDYWGQISVTVS 120
QY 116 S 116
Db 121 S 121
```

Search completed: October 7, 2003, 19:20:00
Job time : 116.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:26 ; Search time 6.2931 Seconds
(without alignments)
126.112 Million cell updates/sec

Title: US-09-988-013A-4_COPY_31_35

Perfect score: 34
Sequence: 1 SWVLH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1880.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1881.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1882.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1883.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1884.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1885.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1886.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1887.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1888.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1889.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1890.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1891.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1892.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1893.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1894.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1895.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1896.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1897.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1898.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1899.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	90	23	ABBS3833
2	34	100.0	96	22	AAU22660
3	34	100.0	96	22	AAW2666
4	34	100.0	112	21	AAV59305
5	34	100.0	116	17	AAW2218
6	34	100.0	116	17	AAW2218
7	34	100.0	116	17	AAW2218
8	34	100.0	116	17	AAW2218
9	34	100.0	116	18	AAW27696
					Variable heavy cha
					Variable heavy cha

10	34	100.0	230	18	AAW07529
11	34	100.0	482	23	ABW48149
12	34	100.0	951	21	AAW44993
13	32	94.1	5	16	AAW74929
14	32	94.1	5	19	AAW70923
15	32	94.1	5	21	AAW19754
16	32	94.1	5	21	AAW92158
17	32	94.1	5	23	AAW11183
18	32	94.1	8	14	AAW35020
19	32	94.1	10	22	AAW35295
20	32	94.1	61	22	AAW63721
21	32	94.1	83	23	ABW63879
22	32	94.1	98	21	AAW40136
23	32	94.1	98	21	AAW40137
24	32	94.1	98	21	AAW40139
25	32	94.1	98	23	ABW91900
26	32	94.1	98	23	ABW91925
27	32	94.1	98	23	ABW78209
28	32	94.1	98	23	ABW78209
29	32	94.1	106	17	AAW91365
30	32	94.1	110	19	AAW70957
31	32	94.1	111	18	AAW04586
32	32	94.1	111	18	AAW04588
33	32	94.1	111	18	AAW04589
34	32	94.1	112	12	AAW12272
35	32	94.1	113	20	AAW95483
36	32	94.1	114	13	AAW21279
37	32	94.1	117	13	AAW30012
38	32	94.1	117	14	AAW30032
39	32	94.1	118	18	AAW27122
40	32	94.1	118	21	AAW56681
41	32	94.1	119	14	AAW39464
42	32	94.1	119	16	AAW79861
43	32	94.1	119	16	AAW79861
44	32	94.1	119	16	AAW79861
45	32	94.1	119	17	AAW04332

ALIGNMENTS

RESULT 1
ABBS3833 standard; Protein: 90 AA.

AC ABB53833;
DE 16-MAY-2002 (first entry)
XX Lactococcus lactis protein yfBg.
DE Lactococcus lactis protein yfBg.
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
OS Lactococcus lactis IL1403.
XX
XX FR2807446-A1.
XX PD 12-OCT-2001.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Bojoline A, Sorokline A, Renault P, Ehrlich SD;
XX WPI: 2002-043418/06.
XX
XX New nucleotide sequence useful in the identification of Lactococcus
XX lactis and related species -
XX Claim 6; SEQ ID NO 535; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (AB90521) and related proteins (AB53300-AB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 90 AA:
Query Match 100.0%; Score 34; DB 23; Length 90;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWLH 5
DB 17 SYWLH 21
RESULT 2
AAU22660
ID AAU22660 standard; Protein; 96 AA.
XX
AC AAU22660;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human colon associated polypeptide #193.
XX
XX Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW cytosstatic.
XX
OS Homo sapiens.
XX
PN WO200155302-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01240.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0230439.
PR 06-SEP-2000; 2000US-023142.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234957.
PR 25-SEP-2000; 2000US-0234958.
PR 25-SEP-2000; 2000US-0234959.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.

XX 11-MAY-1998; 98EP-0108534.
PR (MICR-) MICROMET GMBH.
XX
PA
XX
PI Rieber EP;
XX WPI; 2000-062300/05.
DR N-PSDB; AA248614.
XX
XX
PT New antibodies against dendritic cells, used to produce products for
PT treating, e.g. cancer -
XX
XX
PS Claim 8; Fig 9; 121pp; English.
XX
XX This sequence represents the DC8 antibody heavy chain variable region
CC (DC8-VH). The invention relates to an antibody that reacts with an
CC epitope on dendritic cells (DCs) displaying features of immature and/or
CC mature DCs from peripheral blood mononuclear cells (PBMCs), but does not
CC react with other PBMCs. The antibodies can be used for identifying or
CC isolating DCs. The DCs can be used for preparing activated
CC antigen-specific T-cells, for identifying an antigen recognisable by
CC T-cells, for identifying T-cell activating or co-stimulation compounds,
CC for identifying compounds which suppress T-cell activation or stimulation
CC and for identifying molecules synthesised by DCs having enhancing
CC modulating or suppressing effect on the antigen-specific activation of
CC T-cells. The products can be used for e.g. treating cancers,
CC infectious diseases, or autoimmune disease. The products can also be
CC used for diagnosis and for production of transgenic animals.
XX
SQ Sequence 112 AA;
XX
XX
Query Match 100.0%; Score 34; DB 21; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYWLH 5
IIIII
DB 31 SYWLH 35
XX
XX
RESULT 5
AAR92216
ID AAR92216 standard; Protein; 116 AA.
XX
AC AAR92216;
XX
XX 28-MAY-1996 (first entry)
DT
XX
DE 28-MAY-1996 (first entry)
XX
XX LL2 Mab VH region.
XX
XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
KW leukaemia; therapy; diagnosis; complementarity determining region;
KW CDR; antibody engineering.
XX
XX Mus musculus.
OS
XX
XX
FH Key Location/Qualifiers
FH Region 31..35
FT /label= CDR1
FT /note= "claim 9, page 44"
FT 50..66
FT /label= CDR2
FT /note= "claim 10, page 45"
FT 99..105
FT /label= CDR3
FT /note= "claim 11, page 45"
XX
XX WO9604925-A1.
PN
XX
XX 22-FEB-1996.
PD
XX
XX 11-AUG-1995; 95WO-US09641.
PF

XX 12-AUG-1994; 94US-0289576.
PR (IMMU-) IMMUNOMEDICS INC.
XX
XX
XX
PI Hansen H, Leung S;
XX WPI; 1996-139454/14.
DR N-PSDB; AAT15802.
XX
XX
PT Chimeric and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell lymphoma(s) and
PT leukaemia(s).
XX
XX
PS Claim 5; Page 36-37; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). These were subsequently linked, respectively, to human
CC kappa and IgG1 constant regions. A humanised Mab was obtd. that
CC retained the B-lymphoma and leukaemia cell targeting and
CC internalisation characteristics of the parental LL2 Mab, and which
CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
CC cytostatic agent for therapeutic applin.
XX
SQ Sequence 116 AA;
XX
XX
Query Match 100.0%; Score 34; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYWLH 5
IIIII
DB 31 SYWLH 35
XX
XX
RESULT 6
AAR92218
ID AAR92218 standard; Protein; 116 AA.
XX
AC AAR92218;
XX
XX 28-MAY-1996 (first entry)
DT
XX
XX Humanised LL2 Mab VH region.
DE
XX
XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
KW leukaemia; therapy; diagnosis; complementarity determining region;
KW CDR; antibody engineering.
XX
XX Synthetic.
OS
XX
XX
FH Key Location/Qualifiers
FH Region 31..35
FT /label= CDR1
FT 50..66
FT /label= CDR2
FT 99..105
FT /label= CDR3
XX
XX WO9604925-A1.
PN
XX
XX 22-FEB-1996.
PD
XX
XX 11-AUG-1995; 95WO-US09641.
PF
XX
XX 12-AUG-1994; 94US-0289576.
PD
XX
XX (IMMU-) IMMUNOMEDICS INC.
PA
XX
XX Hansen H, Leung S;
PI

```

XX WPI; 1996-139454/14.
DR N-PSDB; AAT15804.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell Lymphoma(s) and
PT Leukemia(s).
XX
XX Claim 5; Page 39; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH.
CC (AAR92218). These were subsequently linked, respectively, to human
CC kappa and IgG1 constant regions. A humanised Mab was obtd. that
CC retained the B-lymphoma and leukaemia cell targeting and
CC internalisation characteristics of the parental LL2 Mab, and which
CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
CC cytostatic agent for therapeutic appln.
XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 34; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
   11111
Db 31 SYWLH 35

RESULT 7
AAR92219
ID AAR92219 standard; Protein; 116 AA.
XX
XX AAR92219;
XX
XX 28-MAY-1996 (first entry)
XX
XX Humanised LL2 Mab VH region (version hLL2-1).
XX
XX Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
KW leukemia; therapy; diagnosis; complementarity determining region;
KM CDR; antibody engineering.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 31..35
FT Region /label= CDR1
FT Region 50..66
FT Region /label= CDR2
FT Region 99..105
FT Region /label= CDR3
XX
XX WO9604925-A1.
XX
XX 22-FEB-1996.
XX
XX 11-AUG-1995; 95WO-US09641.
XX
XX 12-AUG-1994; 94US-0289576.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen H, Leung S;
XX
XX WPI; 1996-139454/14.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
PT for the therapy and diagnosis of B-cell Lymphoma(s) and
PT Leukemia(s).
PT

```

```

XX Example 1; Page 40; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). In an alternative version, a glutamine was introduced
CC at position 5 of the humanised VH (AAR92219) to include a PstI site
CC useful for subcloning. The humanised VK and VH were subsequently
CC linked, respectively, to human kappa and IgG1 constant regions. A
CC humanised Mab was obtd. that retained the B-lymphoma and leukaemia
CC cell targeting and internalisation characteristics of the parental
CC LL2 Mab, and which exhibited a lowered HAMA reaction. It can be
CC linked to a cytostatic agent for therapeutic appln.
XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 34; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
   11111
Db 31 SYWLH 35

RESULT 8
AAW27696
ID AAW27696 standard; Protein; 116 AA.
XX
XX AAW27696;
XX
XX 14-APR-1998 (first entry)
XX
XX Variable heavy chain of Mab LL2.
XX
XX Variable heavy chain; B cell; monoclonal antibody; Mab; LL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine;
KM diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
XX Mus sp.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 31..35
FT Region /note="complementarity determining region 1"
FT Region 50..66
FT Region /note="complementarity determining region 2"
FT Region 99..105
FT Region /note="complementarity determining region 3"
XX
XX WO9734632-A1.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04196.
XX
XX 20-MAR-1996; 96US-0013709.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen H, Leung S, Qu Z;
XX
XX WPI; 1997-479995/44.
XX
XX N-PSDB; AAT86129.
XX
XX Monoclonal antibody engineered to contain glycosylation site - in
PT non-Fc constant heavy or light chain region, useful to diagnose or
PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
XX Example 3; Fig 4B; 86pp; English.
XX

```

XX The present sequence is the variable heavy chain of the
CC B cell specific monoclonal antibody (Mab) IL2, which contains an
CC engineered tri-peptide N-glycan acceptor sequence. IL2 is a highly
CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
CC murine Mab. The Mab can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia. The glycosylation site allows a label or
CC therapeutic agent of increased size to be conjugated to the
CC carbohydrate moiety, without affecting the Mab's binding affinity
CC or specificity.

XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 34; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|||||
DB 31 SYWLH 35

RESULT 9
AAW27698
ID AAW27698 standard; Protein: 116 AA.
XX
AC AAW27698;
XX
DT 14-APR-1998 (first entry)
XX
DE Variable heavy chain of Mab hLL2.
XX
KM Variable heavy chain; B cell; monoclonal antibody; Mab; hLL2;
KM B cell lymphoma; lymphocytic leukaemia cell; murine; humanised;
KM diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KM chronic lymphocytic leukaemia.
XX
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 31..35
FT /note= "complementarity determining region 1"
FT Region 50..66
FT /note= "complementarity determining region 2"
FT Region 99..105
FT /note= "complementarity determining region 3"
XX
PN WO9734632-A1.
PD 25-SEP-1997.
XX
PF 19-MAR-1997; 97WO-US04196.
XX
PR 20-MAR-1996; 96US-0013709.
XX
PA (IMMU-) IMMUNOMEDICS INC.
XX
PI Hansen H, Leung S, Qu Z;
XX WPI: 1997-479995/44.
XX DR N-PSDB: AAT88131.
XX
PT Monoclonal antibody engineered to contain glycosylation site - in
PT non-Fc constant heavy or light chain region, useful to diagnose or
PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
PS Example 3; Fig 5B; 88pp; English.
XX
CC The present sequence is the variable heavy chain of the
CC B cell specific monoclonal antibody (Mab) hLL2. hLL2 is a highly
CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell

CC humanised murine Mab. The Mab can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia.
XX

XX
SQ Sequence 116 AA;

Query Match 100.0%; Score 34; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|||||
DB 31 SYWLH 35

RESULT 10
AAW07529
ID AAW07529 standard; Protein: 230 AA.
XX
AC AAW07529;
XX
DT 04-MAR-1997 (first entry)
XX
DE Anti-HGF receptor Mab 5D5 Fab heavy chain.
XX
KM Hepatocyte growth factor receptor; HGF; c-Met; antagonist;
KM monoclonal antibody; Mab; Fab; breast cancer; pancreas cancer;
KM colon cancer; lung cancer; therapy.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT MISC-difference 1
FT /label= Glu, Gln
XX
PN WO9638557-A1.
XX
PD 05-DEC-1996.
XX
PF 31-MAY-1996; 96WO-US08094.
XX
PR 02-JUN-1995; 95US-0460368.
XX
PA (GETH) GENENTECH INC.
XX
PI Schwall RH, Tabor KH;
XX
DR WPI: 1997-034371/03.
DR N-PSDB: AAT43874.
XX
PT Hepatocyte growth factor receptor antagonist - specifically
PT anti-c-Met receptor monoclonal antibody 5D5 Fab, useful for
PT mammalian cancer treatment
XX
PS Claim 12; Fig 1B; 69pp; English.
XX
CC 5D5 Fab comprises a light chain (AAW07528) and heavy chain (AAW07529)
CC of monoclonal antibody (Mab) 5D5, a murine Mab raised against the
CC hepatocyte growth factor (HGF) receptor (c-Met). The Fab can be
CC obtd. by papain digestion of the Mab or by recombinant methods
CC utilising cDNA clones (AAT43873-74) coding for the light and heavy
CC chains. 5D5 Fab is an antagonist of the HGF receptor and is useful
CC for the diagnosis and treatment of breast, pancreas, colon and lung
CC cancer.
XX
SQ Sequence 230 AA;

Query Match 100.0%; Score 34; DB 18; Length 230;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|||||

Db			31 SYWLH 35
RESULT 11			
ID	ABBA8149		
XX	ABBA8149 standard; Protein; 482 AA.		
AC	ABBA8149;		
XX			
DT	05-FEB-2002 (first entry)		
XX			
DE	Listeria monocytogenes protein #853.		
XX			
KW	Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;		
KW	vitamin B12; bacterial infection; disease.		
XX			
OS	Listeria monocytogenes .		
PN	WO200177335-A2.		
PD			
PE	18-OCT-2001.		
PF			
PR	11-APR-2001; 2001WO-FR01118.		
XX			
PR	11-APR-2000; 2000FR-0004629.		
PA	(INSP) INST PASTEUR.		
PI	Buchrieser C, Frangeul L, Couve E, Rusniok C, Fahi H, Dehoux P;		
PI	Dushorget O, Chetoui F, Nedjari H, Glaeser P, Kunst F, Coscart P;		
PI	Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;		
PI	Domínguez-Bernal G, Garrido-García P, Pérez-Martínez A, Amend A;		
PI	Charribort T, Doman E, Haln T, Berche P, Charbit A, Durant L;		
PI	Pérez-Díaz J, Baquero F, García Del Portillo F, Gómez-López N;		
PI	Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;		
DR	Rose M, Voss H;		
XX	WPt: 2002-010914/01.		
PT	Genomic sequence for Listeria monocytogenes, useful e.g. for treatment		
PT	and prevention of Listeria and related bacterial infections, and		
PT	related polypeptides -		
PS	Claim 6; SEQ ID No 854; 192pp; French.		
CC			
CC	The present invention relates to the genome sequence of Listeria		
CC	monocytogenes EGD-e (see ABM03041). The genome sequence and fragments of		
CC	it are useful for selecting probes and primers for detecting genes in L.		
CC	monocytogenes and related organisms, and for studying genetic		
CC	polymorphisms and other genomes. The present sequence is a protein		
CC	encoded by the genome sequence of the present invention. Proteins		
CC	expressed from the genome sequence are useful for raising specific		
CC	antibodies. Identification of L. monocytogenes and related organisms, and		
CC	for biosynthesis and biodegradation, especially biosynthesis of Vitamin		
CC	B12. The genome sequence and proteins encoded by it are also useful for		
CC	selecting compounds that regulate gene expression and cell replication		
CC	and modulate L. monocytogenes-related diseases. In addition, the genome		
CC	sequence and proteins encoded by it are useful in pharmaceutical and		
CC	vaccines compositions for the treatment or prevention of infections by L.		
CC	monocytogenes and related organisms.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
SQ	Sequence 482 AA:		
Query Match	100.0%; Score 34; DB 23; Length 482;		
Best Local Similarity	100.0%; Pred. No. 4.4e+02;		
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0		
OY	1 SYWLH 5		
DB	314 SYWLH 318		

RESULT 12	
AA444993	
ID	AA444993 standard; Protein; 951 AA.
AC	
XX	AA444993;
XX	
DT	23-MAY-2000 (first entry)
XX	
DE	DC8scfv-erbB2EC fusion protein containing tetramerisation domain.
XX	
KW	DC8 scfv; single-chain variable fragment; erbB2EC; extracellular domain;
KW	human; fusion construct; tetramerisation domain; constant domain;
KW	heteroinhibitory; multifunctional compound;
KW	immunoglobulin; cytosolic; immunostimulatory; antileukemia; diagnosis;
KW	antiproliferative; prevention; treatment; malignant; haematopoietic cell;
KW	lymphoma; leukemia; solid tumour; carcinoma; melanoma; sarcoma.
XX	
OS	Chimeric - Unidentified.
OS	Chimeric - Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	1..19
FT	/label= Leader_sequence
FT	20..127
FT	/label= DC8scfv_light_chain_variable_region
FT	128..142
FT	/label= Glycine-Serine-linker
FT	143..254
FT	/label= DC8scfv_heavy_chain_variable_region
FT	255..266
FT	/note= '5' end of human IgG3 upper hinge region
FT	with additional residues"
FT	267..305
FT	/label= Human_p53_tetramerisation_domain
FT	306..312
FT	/label= Short_peptide_linker
FT	313..945
FT	/label= erbB2EC_domain
FT	946..951
FT	/label= His_tag
XX	
PN	WO200006605-A2.
XX	
PD	10-FEB-2000.
XX	
PF	28-JUL-1999; 99MO-EP05416.
XX	
PR	28-JUL-1998; 98EP-0114082.
XX	
PA	(MCCR-) MICROMET GRS BIOMEDIZINISCHE FORSCHUNG.
XX	
PI	Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;
XX	
DR	WPI: 2000-195265/17.
DR	N-PSDB; AA250586.
XX	
PT	New multifunctional compounds useful for preventing and/or treating
PT	malignant cell growth and for detection and diagnosis -
XX	
PS	Example 9; Fig 49; 166pp; English.
XX	
CC	The patent discloses heteroinhibitors which are multifunctional compounds
CC	produced in a mammalian host cell as a secretable and fully functional
CC	heterodimer of two polypeptide chains, where one of the polypeptide
CC	chains comprises, a CH1-domain (constant domain of an immunoglobulin
CC	heavy chain) and the other chain comprises CH2-domain (constant domain of
CC	an immunoglobulin light chain). The polypeptide chains further comprise,
CC	fused to the constant domains at least two (poly)peptides having
CC	different receptor or ligand functions, where further at least two of the
CC	different (poly)peptides lack an intrinsic affinity for one another and
CC	are linked via the constant domains. The heteroinhibitors have

CC cytostatic, immunostimulatory, antileukaemia and antiproliferative
CC activities. These compounds can be used for diagnosing, preventing and
CC treating malignant cell growth related to malignancies of haematopoietic
CC cells e.g. lymphomas and leukaemias, or to solid tumours e.g. carcinomas,
CC melanomas and sarcomas.
CC The present sequence is a fusion protein comprising DCR
CC single-chain Fv (scFv) fragment at the N-terminus, extracellular
CC domain of human erbB2 at the C-terminus and a tetramerisation
CC domain between them. This construct was prepared to find out whether
CC an oligomerisation domain characterised in bacterial expression system
CC is applicable for expression of fully functional and secretable
CC recombinant protein in mammalian host cells. This tetrameric construct
CC was not expressed as secretable and fully functional protein
CC in mammalian cells. Hence general applicability of the tetramerisation
CC domain for oligomerisation strategies in mammalian cells was ruled out.
CC
XX
SQ Sequence 951 AA;

Query Match 100.0%; Score 34; DB 21; Length 951;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|:|:|:
DB 173 SYWLH 177

RESULT 13
AAW74929
ID AAW74929 standard; peptide: 5 AA.
AC AAW74929;
XX
DT 19-JAN-1996 (first entry)
XX
DE H-CDR-1 of anti-idiotype antibody against human anticancer antibody.
KM Anticancer; cancer; CDR; heavy chain; light chain; immunoglobulin;
KW Complementarity determining region.
XX
OS Mus sp.
XX
PN JF07101999-A.
XX
PD 18-APR-1995.
XX
PF 06-OCT-1993; 93JP-0272950.
XX
PR 06-OCT-1993; 93JP-0272950.
PA (HAGI/) HAGIWARA Y.
XX
DR WPI; 1995-182987/24.
XX
PT Novel anti-idiotype antibody against an human anticancer monoclonal
PT antibody - and DNA sequences encoding the antibody, useful in
PT pharmacology, medicine and biochemical fields.
XX
PS Claim 1; Page 2; 28pp; Japanese.
XX
CC A new anti-idiotype antibody against a human anticancer monoclonal
CC antibody is claimed. This antibody contains in its heavy chain 3
CC complementarity determining regions CDR1 (AAW74929-R74931), CDR2
CC (AAW74932-R74935) and CDR3 (AAW74936-R74939), this is also true of the
CC light chain which has its own CDR1 (AAW74944-R74946 and AAW85774), CDR2
CC (AAW74947-R74949) and CDR3 (AAW74950-R74954). The antibody and DNA
CC encoding it are useful in pharmacological, medical and biochemical
CC fields.
XX
SQ Sequence 5 AA;

Query Match 94.1%; Score 32; DB 16; Length 5;
Best Local Similarity 80.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|:|:|:
DB 1 SYWLH 5

RESULT 14
AAW70923
ID AAW70923 standard; peptide: 5 AA.
XX
AC AAW70923;
XX
DT 18-FEB-1999 (first entry)
XX
DE CDR1 of the heavy chain of monoclonal antibody RV-133.
XX
KM Pathogenic virus; tropism: mucosa; CDR region; monoclonal antibody;
KM respiratory syncytial virus; RSV; VP6 protein; rota virus; RV;
KM viral infection; inhibit; fusion; protection; transcription;
KM antiviral agent; prophylaxis; diagnosis; infection; contamination.
XX
OS Synthetic.
OS Mus sp.
XX
PN FR2758331-A1.
XX
PD 17-JUL-1998.
XX
PF 14-JAN-1997; 97FR-0000300.
XX
PR 14-JAN-1997; 97FR-0000300.
XX
PA (UYBO-) UNIV BOURGOGNE.
XX
PI Bourgeois C, Kohl E, Pothier P;
XX
DR WPI; 1998-390320/34.
XX
PT New peptide(s) recognising viral epitope with tropism to mucosa -
PT useful for, e.g. diagnosing, preventing and treating viral
PT infection(s)
XX
PS Claim 12; Page 30; 51pp; French.
XX
CC AAW70905-46 represent peptide sequences that can recognise, by
CC antigen-antibody type reactions, at least 1 epitope of a pathogenic
CC virus having tropism for the mucosa. AAW70905-16 and AAW70929-46 are
CC analogous to CDR regions of monoclonal antibodies specific for
CC respiratory syncytial virus (RSV). AAW70917-28 are analogous to CDR
CC regions of monoclonal antibodies specific for site III or IV of the
CC VP6 protein of rota virus (RV). The peptides can neutralise viral
CC infections and may also inhibit fusion between infected and uninfected
CC cells or cells and viruses. They provide passive or active protection
CC and/or inhibit transcription of the virus, so are useful as antiviral
CC agents or for prophylaxis, in human or veterinary medicine. The peptides
CC can be labelled and used to diagnose infection or contamination by the
CC virus. The peptides are particularly directed against RSV or RS but may
CC also be used against papilloma, adeno, entero, polio, influenza or
CC immune deficiency viruses.
XX
SQ Sequence 5 AA;

Query Match 94.1%; Score 32; DB 19; Length 5;
Best Local Similarity 80.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
|:|:|:
DB 1 SYWLH 5

RESULT 15

```

AAB19754
ID AAB19754 standard; Peptide; 5 AA.
XX
AC AAB19754;
XX
DT 19-FEB-2001 (first entry)
XX
DE Erythropoietin receptor agonist MAb 3G9 VH region CDR1.
XX
KM Erythropoietin receptor; agonist; monoclonal antibody; MAb; mouse;
XX heavy chain variable region; antibody engineering; erythropoiesis;
XX anemia; cytopenia; acute renal failure; antianemic; therapy;
XX complementarity determining region; CDR.
XX
OS Mus sp.
XX
PN WO200061637-A1.
XX
PD 19-OCT-2000.
XX
PF 14-APR-2000; 2000WO-US10284.
XX
PR 14-APR-1999; 99US-0129263.
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
XX
PI Erickson-Miller CL, Holmes SD, Taylor AH, Young PR;
XX
DR WPI; 2000-679469/66.
XX
PT Novel erythropoietin receptor agonist antibody useful for enhancing
XX erythropoiesis in the treatment of anemia, cytopenia or acute renal
XX failure
XX
PS Claim 28; Page 50; 70pp; English.
XX
CC The present sequence is that of complementarity determining region
CC 1 (CDR1) of the heavy chain variable region (VH) (see AAB19745) of
CC 3G9, a murine erythropoietin receptor (Epor) agonist monoclonal
CC antibody. Claimed Epor agonist antibodies comprise a VH region
CC that includes the 3G9 VH CDRs. These include humanised agonist
CC antibodies in which the 3G9 VH CDRs are incorporated into a human
CC framework. The Epor agonist antibodies are used in a claimed
CC method for enhancing erythropoiesis, for the treatment of anemia,
CC cytopenia, acute renal failure, and other conditions with depressed
CC erythrocyte production. Agonist antibodies of Epor have the same
CC therapeutic utility as the natural ligand, but with the advantages
CC of easier purification and longer half-life in vivo.
XX
SQ Sequence 5 AA;
XX
Query Match 94.1%; Score 32; DB 21; Length 5;
Best Local Similarity 80.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWLH 5
1111
DB 1 SYWMH 5

```

Search completed: October 7, 2003, 19:13:57
 Job time : 8.2931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 2.58621 Seconds
(without alignments)
305.878 Million cell updates/sec

Title: US-09-988-013a-4_COPY_31_35
Perfect score: 34
Sequence: 1 SYWLH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	34	100.0	96	11	US-09-764-872-437
2	34	100.0	116	10	US-09-741-843-4
3	34	100.0	116	10	US-09-741-843-8
4	34	100.0	116	10	US-09-741-843-9
5	34	100.0	116	11	US-09-894-839-4
6	34	100.0	116	11	US-09-894-839-8
7	34	100.0	116	11	US-09-894-839-21
8	34	100.0	116	11	US-09-888-013A-4
9	34	100.0	116	11	US-09-988-013A-8
10	34	100.0	230	10	US-09-988-013A-9
11	34	100.0	230	10	US-09-995-693-2
12	34	100.0	230	15	US-10-232-408-2
13	32	94.1	5	9	US-09-748-960-12
14	32	94.1	32	15	US-10-233-996-53
15	32	94.1	83	9	US-09-867-550-498

16	32	94.1	98	15	US-10-194-975-33	Sequence 33, Appl
17	32	94.1	118	9	US-09-905-243-78	Sequence 78, Appl
18	32	94.1	119	9	US-09-905-243-79	Sequence 79, Appl
19	32	94.1	119	15	US-10-233-996-2	Sequence 2, Appl
20	32	94.1	119	15	US-10-233-996-4	Sequence 4, Appl
21	32	94.1	119	15	US-10-233-996-39	Sequence 39, Appl
22	32	94.1	120	15	US-10-096-246-12	Sequence 12, Appl
23	32	94.1	122	8	US-08-779-784-28	Sequence 28, Appl
24	32	94.1	122	15	US-10-096-246-10	Sequence 10, Appl
25	32	94.1	122	15	US-10-096-246-11	Sequence 11, Appl
26	32	94.1	140	9	US-09-748-960-4	Sequence 4, Appl
27	32	94.1	142	9	US-09-564-329A-13	Sequence 13, Appl
28	32	94.1	142	10	US-09-855-153-13	Sequence 13, Appl
29	32	94.1	142	10	US-09-854-811-13	Sequence 13, Appl
30	32	94.1	142	10	US-09-934-773-13	Sequence 13, Appl
31	32	94.1	142	10	US-09-963-620-13	Sequence 13, Appl
32	32	94.1	142	11	US-09-853-632-13	Sequence 13, Appl
33	32	94.1	142	12	US-10-224-720-13	Sequence 13, Appl
34	32	94.1	142	12	US-10-225-779-13	Sequence 13, Appl
35	32	94.1	142	15	US-10-225-784-13	Sequence 13, Appl
36	32	94.1	143	8	US-08-779-784-26	Sequence 26, Appl
37	32	94.1	180	9	US-09-748-960-6	Sequence 6, Appl
38	32	94.1	242	11	US-09-880-748-2084	Sequence 2084, Ap
39	32	94.1	331	15	US-10-059-261-169	Sequence 169, Ap
40	31	91.2	177	11	US-09-768-235B-20	Sequence 20, Appl
41	31	91.2	383	11	US-09-978-418-38	Sequence 38, Appl
42	31	91.2	384	11	US-09-764-891-4366	Sequence 4366, Ap
43	31	91.2	416	9	US-09-925-301-1046	Sequence 1046, Ap
44	30	88.2	61	15	US-10-083-357-856	Sequence 856, Appl
45	30	88.2	132	14	US-10-056-359-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-764-872-437
; Sequence 437, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 437
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-872-437

Query Match
Best Local Similarity 100.0%; Score 34; DB 11; Length 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
DB 32 SYWLH 36

RESULT 2
US-09-741-843-4
; Sequence 4, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
; TITLE OF INVENTION: AND LEUKEMIA CELLS
; FILE REFERENCE: 018733/0996

```

; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-741-843-4

Query Match
Best Local Similarity 100.0%; Score 34; DB 10; Length 116;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
Db 31 SYWLH 35

RESULT 3
US-09-741-843-8
; Sequence 8, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-8

Query Match
Best Local Similarity 100.0%; Score 34; DB 10; Length 116;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
Db 31 SYWLH 35

RESULT 4
US-09-741-843-9
; Sequence 9, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
```

```

; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-9

Query Match
Best Local Similarity 100.0%; Score 34; DB 10; Length 116;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
Db 31 SYWLH 35

RESULT 5
US-09-894-839-4
; Sequence 4, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-4

Query Match
Best Local Similarity 100.0%; Score 34; DB 11; Length 116;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
Db 31 SYWLH 35

RESULT 6
US-09-894-839-8
; Sequence 8, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
```

```

; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-8

Query Match          100.0%; Score 34; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
        |||||
Db      31 SYWLH 35

RESULT 7
US-09-894-839-21
; Sequence 21, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-21

Query Match          100.0%; Score 34; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
        |||||
Db      31 SYWLH 35

RESULT 8
US-09-988-013A-4
; Sequence 4, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
```

```

; ORGANISM: Murinae gen. sp.
US-09-988-013A-4

Query Match          100.0%; Score 34; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
        |||||
Db      31 SYWLH 35

RESULT 9
US-09-988-013A-8
; Sequence 8, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-988-013A-8

Query Match          100.0%; Score 34; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
        |||||
Db      31 SYWLH 35

RESULT 10
US-09-988-013A-9
; Sequence 9, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
```

TYPE: PRT
ORGANISM: Homo sapiens
US-09-988-013a-9

Query Match 100.0%; Score 34; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
11111
DB 31 SYWLH 35

RESULT 11
US-09-995-693-2
Sequence 2, Application US/09995693
Patent No. US20020136721A1
GENERAL INFORMATION:
APPLICANT: Schwall, Ralph H.
Tabor, Kelly H.

TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
Antagonists and Uses Thereof

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPacIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/995,693

FILING DATE: 29-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/952,235

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Marschang, Diane L.

REGISTRATION NUMBER: 35,600

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-5416

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 230 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 100.0%; Score 34; DB 10; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
11111
DB 31 SYWLH 35

RESULT 12
US-10-232-408-2

Sequence 2, Application US/10232408

Publication No. US20030118587A1

GENERAL INFORMATION:

APPLICANT: Schwall, Ralph H.
Tabor, Kelly H.

LOCATION: (1)...(5)

TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
Antagonists and Uses Thereof

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPacIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/232,408

FILING DATE: 03-Sep-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/669,971

FILING DATE: 05-Jul-2001

APPLICATION NUMBER: US/08/952,235

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/460368

FILING DATE: 02-Jun-1995

ATTORNEY/AGENT INFORMATION:

NAME: Marschang, Diane L.

REGISTRATION NUMBER: 35,600

REFERENCE/DOCKET NUMBER: P0938P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-5416

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 230 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 100.0%; Score 34; DB 15; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
11111
DB 31 SYWLH 35

RESULT 13
US-09-748-960-12

Sequence 12, Application US/09748960

Patent No. US20010046496A1

GENERAL INFORMATION:

APPLICANT: Bretzman, Lee R.

APPLICANT: Allison, David Edward

TITLE OF INVENTION: Method of Administering an Antibody

FILE REFERENCE: 1855.2007-001

CURRENT APPLICATION NUMBER: US/09/748,960

PRIOR FILING DATE: 2000-12-27

PRIOR APPLICATION NUMBER: US 09/550,082

PRIOR FILING DATE: 2000-04-14

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12

LENGTH: 5

TYPE: PRT

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: SITE

LOCATION: (1)...(5)

```
; OTHER INFORMATION: CDRL of the heavy chain of antibodies Act-1 and
; OTHER INFORMATION: LDP-02
US-09-748-960-12

Query Match          94.1%; Score 32; DB 9; Length 5;
Best Local Similarity 80.0%; Pred. No. 5.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
       |||:|
Db      1 SYWMH 5

RESULT 14
US-10-233-996-53
; Sequence 53, Application US/10233996
; Publication No. US20030096976A1
; GENERAL INFORMATION:
; APPLICANT: HONG, Hyo Jeong
; APPLICANT: PARK, Sung Sup
; APPLICANT: KANG, Young Jun
; APPLICANT: KANG, Chang-Yu11
; APPLICANT: YOON, Sung Kwan
; TITLE OF INVENTION: HUMANIZED ANTIBODIES LB-00503 AND LB-00506 SPECIFIC FOR HUMAN 4-1
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING SAID HUMANIZED ANTIBODIES
; FILE REFERENCE: 4363-0102P
; CURRENT APPLICATION NUMBER: US/10/233,996
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 32
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: Humanized antibody heavy chain directed against a human 4-1BB
; OTHER INFORMATION: molecule
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (17)..(18)
; OTHER INFORMATION: X = any amino acid, unknown or other
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (21)..(22)
; OTHER INFORMATION: X = any amino acid, unknown or other
US-10-233-996-53

Query Match          94.1%; Score 32; DB 15; Length 32;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
       |||:|
Db      1 SYWMH 5

RESULT 15
US-09-867-550-498
; Sequence 498, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
```

```
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 498
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (57)
; OTHER INFORMATION: wherein Xaa may be any one of
US-09-867-550-498

Query Match          94.1%; Score 32; DB 9; Length 83;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYWLH 5
       |||:|
Db      69 SYWMH 73

Search completed: October 7, 2003, 19:24:27
Job time : 3.58621 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 1.86207 Seconds
(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013A-4_COPY_31_35

Perfect score: 34

Sequence: 1 SYWLH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	90	2 D86690	hypothetical prote
2	34	100.0	196	2 T01982	tumor related prot
3	34	100.0	317	2 E86264	protein R3F19.7 [i
4	34	100.0	482	2 AFI120	conserved hypotnet
5	34	100.0	494	2 A11480	protein T08G11.4 [
6	34	100.0	495	2 B87869	hypothetical prote
7	34	100.0	566	2 T24696	probable peptide s
8	34	100.0	4924	2 T50176	ig heavy chain v r
9	32	94.1	35	2 S46472	ig heavy chain v r
10	32	94.1	50	2 S26941	ig heavy chain v r
11	32	94.1	56	2 E49038	ig heavy chain v r
12	32	94.1	58	2 D49038	ig lambda chain v
13	32	94.1	69	2 D25150	ig heavy chain v r
14	32	94.1	71	2 PH1136	ig heavy chain v r
15	32	94.1	76	2 PH1153	ig heavy chain v r
16	32	94.1	80	2 F28833	ig heavy chain v r
17	32	94.1	86	2 A25150	ig kappa chain v r
18	32	94.1	90	2 PH1152	ig heavy chain v r
19	32	94.1	94	2 J10078	ig heavy chain v r
20	32	94.1	94	2 PH1142	ig heavy chain v r
21	32	94.1	94	2 S42185	ig gamma chain v r
22	32	94.1	96	2 S17616	ig heavy chain v r
23	32	94.1	97	2 PH1137	ig heavy chain v r
24	32	94.1	98	2 P10121	ig heavy chain v r
25	32	94.1	98	2 PH1128	ig heavy chain v r
26	32	94.1	98	2 PH1129	ig heavy chain v r
27	32	94.1	98	2 PH1131	ig heavy chain v r
28	32	94.1	98	2 PH1134	ig heavy chain v r
29	32	94.1	98	2 PH1138	ig heavy chain v r

30	32	94.1	98	2 PH1139	ig heavy chain v r
31	32	94.1	98	2 PH1141	ig heavy chain v r
32	32	94.1	98	2 PH1147	ig heavy chain v r
33	32	94.1	98	2 PH1149	ig heavy chain v r
34	32	94.1	98	2 PH1150	ig heavy chain v r
35	32	94.1	98	2 PH1154	ig heavy chain v r
36	32	94.1	98	2 PH1156	ig heavy chain v r
37	32	94.1	98	2 PH1157	ig heavy chain v r
38	32	94.1	98	2 PH1164	ig heavy chain v r
39	32	94.1	98	2 PH1105	ig heavy chain v r
40	32	94.1	98	2 PH1108	ig heavy chain v r
41	32	94.1	98	2 PH1111	ig heavy chain v r
42	32	94.1	98	2 PH1114	ig heavy chain v r
43	32	94.1	98	2 PH1118	ig heavy chain v r
44	32	94.1	98	2 PH1119	ig heavy chain v r
45	32	94.1	98	2 PH1122	ig heavy chain v r

ALIGNMENTS

RESULT 1
D86690
hypothetical protein yfbc [imported] - Lactococcus lactis subsp. lactis (strain IL140
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86690
R:Bohlin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: D86690
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <STO>
A:Cross-references: GB:AE005176; PTD:g12723408; PIDN:AAK04622.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yfbc

Query Match 100.0%; Score 34; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
DB 17 SYWLH 21

RESULT 2
T01982
tumor related protein HR4 - (common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 19-May-2000
C:Accession: T01982
R:Nakamura, C.; Nishimura, N.; Mishima, M.; Ueda, T.; Onaka, T.; Tsvetanov, S.; Takum
submitted to the EMBL Data Library, January 1998
A:Description: Molecular cloning and characterization of cDNAs preferentially express
A:Reference number: Z14477
A:Accession: T01982
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-196 <NAK>
A:Cross-references: EMBL:AF043554; NID:g2852376; PTD:g2852377
A:Experimental source: tissue-type hairy root tumor; crown gall tumor
C:Genetics:
A:Gene: HR4
C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 100.0%; Score 34; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 153 SYWLH 157

RESULT 3

E86264
protein F3F19.7 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cross)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: E86264
R/Domiguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlan, K.D.; Fshl, H.
R/Neologs, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
aneen, N.F.; Hughes, B.; Hutzler, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; PMID:21016719; PMID:11130712
A/Accession: E86264
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-317 <STO>
A/Cross-references: GB:AE005172; NID:g4850388; PIDN:AAD31058.1; GSPDB:GN00141
C/Genetics:
A/Gene: F3F19.7
A/Map position: 1

Query Match 100.0%; Score 34; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 289 SYWLH 293

RESULT 4

AF1120
conserved hypothetical protein homolog lmo0365 [imported] - Listeria monocytogenes (str
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: AF1120
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlan, K.D.; Fshl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; PMID:21537279; PMID:11679669
A/Accession: AF1120
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-482 <GLA>
A/Cross-references: GB:NC_003210; PIDN:CAC98444.1; PID:g16409743; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo0365
C/Superfamily: conserved hypothetical protein ywbl

Query Match 100.0%; Score 34; DB 2; Length 482;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 314 SYWLH 318

RESULT 5

Al1480
conserved hypothetical protein homolog lln0384 [imported] - Listeria innocua (strain
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: Al1480
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlan, K.D.; Fshl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.;
ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; PMID:21537279; PMID:11679669
A/Accession: Al1480
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-494 <GLA>
A/Cross-references: GB:AL592022; PIDN:CAC95617.1; PID:g16412813; GSPDB:GN00178
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: lln0384
C/Superfamily: conserved hypothetical protein ywbl

Query Match 100.0%; Score 34; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 314 SYWLH 318

RESULT 6

B87869
protein T08G11.4 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: B87869
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A/Reference number: A75000; PMID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/Celegans/ and www.sanger.ac.uk/Projects/C-
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A/Accession: B87869
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-495 <STO>
A/Cross-references: GB:chr_1; PIDN:CAB02307.1; PID:g3879758; GSPDB:GN00019; CESP:T08G
C/Genetics:
A/Gene: T08G11.4
A/Map position: 1

Query Match 100.0%; Score 34; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 62 SYWLH 66

RESULT 7

T24696
hypothetical protein T08G11.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T24696
R/Dobson, R.
submitted to the EMBL Data Library, September 1996
A/Reference number: Z19925
A/Accession: T24696

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-566 <M12>
A:Cross-references: EMBL:Z80220; PIDN:CA802307.2; GSPDB:GN00019; CESP:T08G11.4
A:Experimental source: clone T08G11
C:Genetics:
A:Gene: CESP:T08G11.4
A:Map position: 1
A:Introns: 47/2; 86/1; 177/3; 235/3; 282/2; 415/1; 521/3

Query Match 100.0%; Score 34; DB 2; Length 566;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 133 SYWLH 137

RESULT 8

T50176
probable peptide synthetase [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 01-Dec-2000
C:Accession: T50176
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25044
A:Accession: T50176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4924 <BAD>
A:Cross-references: EMBL:AL138854; PIDN:CA872227.1; GSPDB:GN00066; SPDB:SPAC23G3.02c
A:Experimental source: strain 972h(-); cosmid c23G3
C:Genetics:
A:Gene: SPDB:SPAC23G3.02c
A:Map position: 1
A:Introns: 2125/1
C:Superfamily: acetate-CoA ligase homology; acyl carrier protein homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:260-697/Domain: acetate-CoA ligase homology <ACLI>
F:2772-3226/Domain: acetate-CoA ligase homology <ACLI2>
F:4405-4474/Domain: acyl carrier protein homology <ACP>
F:2206-3288/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 100.0%; Score 34; DB 2; Length 4924;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 1794 SYWLH 1798

RESULT 9

S46472
Ig heavy chain V region (DA-8) - human
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Jun-2000
C:Accession: S46472
R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Rietman, H.; Carter, N.P.; Bulwela, L.; Wit
Nature Genet. 7, 162-168, 1994
A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telom
A:Reference number: S46460; MUID:95004581; PMID:7920635
A:Accession: S46472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <COO>
A:Cross-references: EMBL:Z30082; NID:9505468; PIDN:CAA82900.1; PID:91335183
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 35;

Best Local Similarity 80.0%; Pred. No. 15;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 4 SYWLH 8

RESULT 10

S26941
Ig heavy chain V region (DP-80) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 17-Nov-1995 #text_change 17-Mar-1999
C:Accession: S26941
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26941
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-50 <TOM>
A:Cross-references: EMBL:Z14076
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 50;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 1 SYWLH 5

RESULT 11

E49038
Ig lambda chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: E49038
R:Weiss, U.; Zobelein, R.; Rajewsky, K.
Eur. J. Immunol. 22, 511-517, 1992
A:Title: Accumulation of somatic mutants in the B cell compartment after primary immu
A:Reference number: A49038; MUID:92164733; PMID:1537385
A:Accession: E49038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-56 <WEI>
A:Cross-references: GB:S85736; NID:g246322; PIDN:AAB21563.1; PID:g246323
A:Experimental source: spleen
A:Note: sequence extracted from NCBI backbone (NCBIN:85736, NCBI:P:85804)
C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 56;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
|||||
Db 10 SYWLH 14

RESULT 12

D49038
Ig lambda chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: D49038
R:Weiss, U.; Zobelein, R.; Rajewsky, K.
Eur. J. Immunol. 22, 511-517, 1992
A:Title: Accumulation of somatic mutants in the B cell compartment after primary immu

A:Reference number: A49038; MUID:92164733; PMID:1537385
 A:Accession: D49038
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-58 <ME1>
 A:Cross-references: GB:S85735; NID:9246320; PIDN:AAB21562.1; PID:9246321
 A:Experimental source: spleen
 A>Note: sequence extracted from NCBI backbone (NCBIN:85735, NCBIPI:85802)
 C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 58;
 Best Local Similarity 80.0%; Pred. No. 24;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
 |||:|
 DB 12 SYWMH 16

RESULT 13

D25150
 Ig heavy chain V region (AC38 260.2) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 16-Aug-1996
 C:Accession: D25150
 R:Didrop, R.; Boyens, J.; Stekevitz, M.; Beyreuther, K.; Rajewsky, K.
 EMBL J. 3, 517-523, 1984
 A:Title: A V region determinant (idiotope) expressed at high frequency in B lymphocytes
 A:Reference number: A91000; MUID:84182519; PMID:6201362
 A:Accession: D25150
 A:Molecule type: protein
 A:Residues: 1-69 <DIL>
 A:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 69;
 Best Local Similarity 80.0%; Pred. No. 28;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
 |||:|
 DB 19 SYWMH 23

RESULT 14

PH1136
 Ig heavy chain V region (clone V2075.1) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 C:Accession: PH1136
 R:Schlitz, B.; Rajewsky, K.
 J. Exp. Med. 176, 427-438, 1992
 A:Title: Natural occurrence and origin of somatically mutated memory B cells in mice.
 A:Reference number: PH1105; MUID:92364545; PMID:1500855
 A:Accession: PH1136
 A:Molecule type: DNA
 A:Residues: 1-71 <SCH>
 A:Experimental source: B cell
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 71;
 Best Local Similarity 80.0%; Pred. No. 29;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
 |||:|
 DB 4 SYWMH 8

RESULT 15

PH1153
 Ig heavy chain V region (clone 43F.1) - mouse (fragment)

C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 C:Accession: PH1153
 R:Schlitz, B.; Rajewsky, K.
 J. Exp. Med. 176, 427-438, 1992
 A:Title: Natural occurrence and origin of somatically mutated memory B cells in mice.
 A:Reference number: PH1105; MUID:92364545; PMID:1500855
 A:Accession: PH1153
 A:Molecule type: DNA
 A:Residues: 1-76 <SCH>
 A:Experimental source: B cell
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 94.1%; Score 32; DB 2; Length 76;
 Best Local Similarity 80.0%; Pred. No. 31;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
 |||:|
 DB 9 SYWMH 13

Search completed: October 7, 2003, 19:21:50
 Job time : 2.86207 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 1.01724 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-4_COPY_31_35.

Perfect score: 34

Sequence: 1 SYWLH 5

Scoring table: BL0SUM62

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	94.1	117	1	HV04_MOUSE
2	32	94.1	117	1	HV06_MOUSE
3	32	94.1	117	1	HV09_MOUSE
4	32	94.1	117	1	HV10_MOUSE
5	32	94.1	117	1	HV49_MOUSE
6	32	94.1	120	1	HV50_MOUSE
7	32	94.1	139	1	HV07_MOUSE
8	32	94.1	449	1	VNSS_INSVN
9	32	94.1	464	1	VNSS_TSMV1
10	32	94.1	467	1	VNSS_TSMV1
11	32	94.1	3080	1	POLG_ZYMYC
12	32	94.1	3083	1	POLG_ZYMYR
13	32	94.1	3083	1	POLG_ZYMYR
14	31	91.2	309	1	YESP_BACSU
15	31	91.2	370	1	TAM2_HUMAN
16	31	91.2	373	1	TAM1_HUMAN
17	31	91.2	466	1	SYC_BACHD
18	31	91.2	473	1	PEPD_IACSK
19	30	88.2	137	1	BFR_NITW1
20	30	88.2	137	1	YH97_MERTH
21	30	88.2	161	1	BFR_BROME
22	30	88.2	166	1	OV22_ONCVO
23	30	88.2	174	1	OV21_ONCVO
24	30	88.2	175	1	VIRR_AGRTR
25	30	88.2	247	1	I125_RICPR
26	30	88.2	264	1	FLIR_SALTY
27	30	88.2	287	1	KISR_BACSU
28	30	88.2	301	1	MTAL_SORFI
29	30	88.2	304	1	TRUB_PSEAE
30	30	88.2	306	1	MTAL_SORMA
31	30	88.2	372	1	B4G2_HUMAN
32	30	88.2	372	1	G391_DROME
33	30	88.2	400	1	SCW_DROME

34	30	88.2	414	1	YAFI_SALTY	P37722 salmonella
35	30	88.2	466	1	SYC_BACSU	Q06752 bacillus su
36	30	88.2	466	1	SYC_CLOPE	Q6xh55 clostridium
37	30	88.2	469	1	SYCI_MYCTU	P36862 mycobacteri
38	30	88.2	473	1	SYCI_MYCLE	P57990 mycobacteri
39	30	88.2	473	1	SYC_NEIMB	Q9jx66 neisseria m
40	30	88.2	474	1	SYC_CHLBN	Q92634 chlamydia p
41	30	88.2	475	1	SYC_AERPE	Q9ybk6 aeropyrum p
42	30	88.2	475	1	SYC_CHLMU	Q9p160 chlamydia m
43	30	88.2	476	1	SYC_PYRHO	Q58370 pyrococcus
44	30	88.2	477	1	SYC_CHLFR	Q84787 chlamydia t
45	30	88.2	477	1	SYC_PYRAB	Q9uyv2 pyrococcus

ALIGNMENTS

RESULT 1					
ID	HV04_MOUSE	STANDARD:	PRT:	117 AA.	
AC	P01748;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	19 heavy chain V region 23 precursor.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_Taxid:10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6;				
RX	MEDLINE=81234548; PubMed=6788376;				
RA	Bothwell A.L.M., Paskind M., Reih M., Imanishi-Karl T., Rajewsky K.,				
RA	Baltimore D.;				
RT	*Heavy chain variable region contribution to the NPb family of				
RT	antibodies: somatic mutation evident in a gamma 2a variable region.;				
RL	Cell 24:625-637(1981).				
CC	-1- MISCELLANEOUS: THIS GERM LINE GENE BELONGS TO A SET OF CLOSELY				
CC	RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.				
CC	PIR: A02030; HWSM23.				
DR	HSSP: P01810; 2PBJ.				
DR	InterPro: IPR007110; Ig-like.				
DR	InterPro: IPR003006; Ig_MHC.				
DR	InterPro: IPR003596; Ig_V.				
DR	Pfam: PF00047; Ig; 1.				
DR	SMART: SM00407; IgV; 1.				
DR	PROSITE: PS50835; IG_LIKE; 1.				
KW	Immunoglobulin V region; Signal.				
FT	SIGNAL	1	19		
FT	CHAIN	20	117		
FT	DOMAIN	50	49		
FT	DOMAIN	50	54		
FT	DOMAIN	55	68		
FT	DOMAIN	69	85		
FT	DOMAIN	86	117		
FT	DISULFID	41	115		
FT	NON_TER	117	117		
SQ	SEQUENCE	117 AA;	12772 MM;	C530F829C906F69B CRC64;	
Query Match					
Best Local Similarity 94.1%; Score 32; DB 1; Length 117;					
Matches 4; Conservativity 1; Pred. No. 22;					
Mismatches 0; Indels 0; Gaps 0;					
Oy	1 SYWLH 5				
Db	50 SYWLH 54				
RESULT 2					
ID	HV06_MOUSE	STANDARD:	PRT:	117 AA.	
AC	P01750;				

```

DR 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 102 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC PIR: A02032; HVMS02.
DR HSP: P01810; 2FBJ.
DR InterPro: IPR007110; Ig_1-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 102.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
SQ SEQUENCE 117 AA; 12867 MW; 740A65DD851FCA8C CRC64;

Query Match 94.1%; Score 32; DB 1; Length 117;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 50 SYWLH 54

RESULT 3
HY09_MOUSE STANDARD; PRT; 117 AA.
AC P01753; P11271;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 186-1 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC PIR: D90809; HVMS61.
DR HSP: P01810; 2FBJ.

```

```

DR InterPro: IPR007110; Ig_1-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 186-1.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
SQ SEQUENCE 117 AA; 12890 MW; 16191A086CB17F5A CRC64;

Query Match 94.1%; Score 32; DB 1; Length 117;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 50 SYWLH 54

RESULT 4
HY10_MOUSE STANDARD; PRT; 117 AA.
AC P01754; P11270;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig heavy chain V region 145 precursor.
GN IGH-VJ558.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J00533; AAA38602.1; -
DR PIR: C90809; HVMS45.
DR HSP: P01810; 2FBJ.
DR MGD: MGI:96486; Igh-VJ558.
DR InterPro: IPR007110; Ig_1-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 145.

```

```
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12921 MW; D37DE8A3F543E996 CRC64;

Query Match 94.1%; Score 32; DB 1; Length 117;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMWH 5
DB 50 SYMWH 54

RESULT 5
HV49_MOUSE
ID HV49_MOUSE STANDARD; PRT; 117 AA.
AC P06328;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH558 B4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "developmentally controlled and tissue-specific expression of
unrearranged VH gene segments."
RL Cell 40:271-281(1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13788; AAA38506.1; -.
DR PIR; A02035; MHMSB4.
DR HSSP; P01810; 2PBJ.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION VH558 B4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12834 MW; B8862FAC67ABD345 CRC64;

Query Match 94.1%; Score 32; DB 1; Length 117;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMWH 5
DB 50 SYMWH 54
```

```
RESULT 6
HV50_MOUSE
ID HV50_MOUSE STANDARD; PRT; 120 AA.
AC P06329;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region AC38 15.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=84182519; PubMed=6201362;
RA Dildrop R., Bovens J., Stekevitz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes."
RL EMBO J. 3:517-523(1984).
DR PIR; A02037; MHMS15.
DR HSSP; P01810; 2PBJ.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 98 D SEGMENT.
FT DOMAIN 99 105 V SEGMENT.
FT DOMAIN 106 120 J SEGMENT.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13311 MW; 914453F426F09834 CRC64;

Query Match 94.1%; Score 32; DB 1; Length 120;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMWH 5
DB 31 SYMWH 35

RESULT 7
HV07_MOUSE
ID HV07_MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region BI-8/186-2 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bochtell A.L.M., Paakind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region."
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THE BI-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC (NPB ANTIBODIES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J00529; AAA38170.1; -

DR PIR: A90809; MHMS18.

DR PDB: 1AGU; 27-MAY-98.

DR PDB: 1AGU; 15-JUL-98.

DR InterPro: IPR007110; Ig-1like.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_v.

DR Pfam: PF00047; Ig_1.

DR SMART: SM00406; IGV; 1.

DR PROSITE: PS50835; IG_LIKE; 1.

KW Immunoglobulin V region; Signal; 3D-structure.

FT SIGNAL 1 19 IG HEAVY CHAIN V REGION BI-8/186-2.

FT CHAIN 20 139 FRAMEWORK-1.

FT DOMAIN 20 49 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 50 54 FRAMEWORK-2.

FT DOMAIN 55 68 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 69 85 FRAMEWORK-3.

FT DOMAIN 86 117 D SEGMENT.

FT DOMAIN 118 134 H2 SEGMENT.

FT DOMAIN 125 139 BY SIMILARITY.

FT DISULFID 41 115

FT NON_TER 139 139

SQ SEQUENCE 139 AA; 15419 MW; 1B57DD4FDC9F465 CRC64;

Query Match Best Local Similarity 94.1%; Score 32; DB 1; Length 139;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5

Db 50 SYWMH 54

RESULT 8 VNSS_INSVN STANDARD; PRT; 449 AA.

ID 001811;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Nonstructural protein NS-S.

GN NSS.

OS Impatiens necrotic spot virus (strain NL-07) (INSV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tosopovirus.

OX NCBI_TaxID-31622;

RN RN SEQUENCE FROM N.A.

RP MEDLINE-92331780; PubMed-1385787;

RA de Haan P., de Avilla A.C., Kormelink R., Westerbroek A.,

RA Gielen J.J., Peters D., Goldbach R.;

RT "The nucleotide sequence of the S RNA of Impatiens necrotic spot

RT virus, a novel tospovirus."

RL FEBS Lett. 306:27-32(1992).

CC -I- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X66972; CAA47382.1; -

DR PIR: S23158; S23158.

DR InterPro: IPR004915; Bunya_NS-S_2.

DR Pfam: PF03231; Bunya_NS-S_2; 1.

KW Nonstructural protein.

SQ SEQUENCE 449 AA; 51197 MW; C46AC1372B11ACAS CRC64;

Query Match Best Local Similarity 94.1%; Score 32; DB 1; Length 449;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5

Db 29 SYWMH 33

RESULT 9 VNSS_TSWVL STANDARD; PRT; 464 AA.

ID VNSS_TSWVL

AC P26002;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Nonstructural protein NS-S.

GN NSS.

OS Tomato spotted wilt virus (strain Brazilian Br-01) (TSWV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tosopovirus.

OX NCBI_TaxID-36413;

RN RN SEQUENCE FROM N.A.

RP MEDLINE-90264829; PubMed-1693160;

RA de Haan P., Wagemakers L., Peters D., Goldbach R.;

RT "The S RNA segment of tomato spotted wilt virus has an ambisense

RT character".

RL J. Gen. Virol. 71:1001-1007(1990).

CC -I- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: D00645; BAA00540.1; -

DR PIR: J00547; MNYTWC.

DR InterPro: IPR004915; Bunya_NS-S_2.

DR Pfam: PF03231; Bunya_NS-S_2; 1.

KW Nonstructural protein.

SQ SEQUENCE 464 AA; 52448 MW; AE5519179F9EF377 CRC64;

Query Match Best Local Similarity 94.1%; Score 32; DB 1; Length 464;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5

Db 29 SYWMH 33

RESULT 10 VNSS_TSWVL STANDARD; PRT; 467 AA.

ID VNSS_TSWVL

AC P26003;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Nonstructural protein NS-S.

GN NSS.

OS Tomato spotted wilt virus (strain Bulgarian L3) (TSWV).

OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tosopovirus.

OX NCBI_TaxID-36415;

RN RN SEQUENCE FROM N.A.

RP MEDLINE-91132150; PubMed-1993884;

RA Maiss E., Ivanova L., Breyel E., Adam G.;

RT Cloning and sequencing of the S RNA from a Bulgarian isolate of
RT tomato spotted wilt virus.";
RT J. Gen. Virol. 72:461-464(1991).
CC -1- FUNCTION: FORMS FILAMENTOUS INCLUSION BODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D13926; BAA03024.1; -.
CC PIR: J00954; MNVWL.
CC DR InterPro: IPR004915; Bunya_NS-S_2.
CC DR Pfam: PF03231; Bunya_NS-S_2; 1.
CC KM Nonstructural protein.
CC SQ SEQUENCE 467 AA; 52413 MW; 385C4AE802DB6DC CRC64;
Query Match 94.1%; Score 32; DB 1; Length 467;
Best Local Similarity 80.0%; Pred. No. 82;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWLH 5
Db 29 SYWLH 33
RESULT 11
POLG_ZYMVC STANDARD; PRT; 3080 AA.
ID POLG_ZYMVC
AC P18479; Q89334;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [contains: N-terminal protein (PI); Helper
DE component proteinase (EC 3.4.22.45) (HC-Pro); protein P3; 6 kDa
DE protein 1 (K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
DE (K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
DE (EC 2.7.7.48); Coat protein (CP)].
OS Zucchini yellow mosaic virus (Strain California) (ZYMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Polyviruses.
OX NCBI_TaxID=117128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95146958; PubMed=7844540;
RA Wisler G.C., Purcell D.E., Hiebert E.;
RT "Characterization of the PI protein and coding region of the zucchini
RT yellow mosaic virus.";
RL J. Gen. Virol. 76:37-45(1995).
RN [2]
RP SEQUENCE OF 2694-3080 FROM N.A.
RX MEDLINE=90236320; PubMed=2185142;
RA Gal-On A., Antignus Y., Rosner A., Raccach B.;
RT "Nucleotide sequence of the zucchini yellow mosaic virus
RT capsid-encoding gene and its expression in Escherichia coli.";
RL Gene 87:273-277(1990).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC further restricted by preferences for glutamyl bonds, and activity is
CC that vary with the species of polyviruses, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC Gln(Ser or Gly) for the enzyme from tobacco etch virus. The
CC natural substrate is the viral polypeptide, but other proteins and
CC oligopeptides containing the appropriate consensus sequence are
CC also cleaved.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC (RNA)(N).
CC -1- CATALYTIC ACTIVITY: Hydrolyzes a gly-l-gly bond at its own C-
CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-l-gly, in the
CC processing of the polyviral polypeptide.
CC -1- PFM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L31350; AAA65559.1; -.
CC PIR: M35095; AAA48511.1; -.
CC DR HSSP: P27958; 1HE1.
CC DR InterPro: IPR001410; DEAD.
CC DR InterPro: IPR001650; Helicase_C.
CC DR InterPro: IPR001730; Peptidase_C4.
CC DR InterPro: IPR001456; Peptidase_C6.
CC DR InterPro: IPR001592; Poly_coat.
CC DR InterPro: IPR002540; Poly_P1.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVLr.
CC Pfmam: PR002271; helicase_C_1.
CC Pfmam: PF00863; Peptidase_C4; 1.
CC Pfmam: PF00851; Peptidase_C6; 1.
CC Pfmam: PF00767; Poly_coat; 1.
CC Pfmam: PF01577; Poly_P1; 1.
CC Pfmam: PF00680; RNA_dep_RNA_pol; 1.
CC PRINTS: PR00966; NTPOTYPIASE.
CC SMART: SM00487; DEXDC; 1.
CC SMART: SM00490; HELICG; 1.
CC KM Hydrolyase; Transferase; Thiol protease; RNA-directed RNA polymerase;
CC KM Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase;
CC KM ATP-binding.
CC FT CHAIN 1 304 N-TERMINAL PROTEIN.
CC FT CHAIN 305 766 HELPER COMPONENT PROTEINASE.
CC FT CHAIN 767 ? PROTEIN P3.
CC FT CHAIN ? 1164 6 kDa PROTEIN 1.
CC FT CHAIN 1165 1798 CYTOPLASMIC INCLUSION PROTEIN.
CC FT CHAIN 1799 1851 6 kDa PROTEIN 2.
CC FT CHAIN 1852 2041 GENOME-LINKED PROTEIN.
CC FT CHAIN 2042 2284 NUCLEAR INCLUSION PROTEIN A.
CC FT CHAIN 2285 2801 NUCLEAR INCLUSION PROTEIN B.
CC FT CHAIN 2802 3080 COAT PROTEIN.
CC FT NP_BIND 1249 1256 ATP (POTENTIAL).
CC FT CONFLICT 2694 2695 LE -> ST (IN REF. 2).
CC FT CONFLICT 2699 2701 IVS -> LFP (IN REF. 2).
CC FT CONFLICT 2811 2811 A -> T (IN REF. 2).
CC FT CONFLICT 2811 2811 G -> S (IN REF. 2).
CC FT CONFLICT 2834 2834
CC SQ SEQUENCE 3080 AA; 350624 MW; 2A1E501DEA6B9F73 CRC64;
Query Match 94.1%; Score 32; DB 1; Length 3080;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWLH 5
Db 2179 SYWLH 2183

```

RESULT 12
POLG_ZYMWV STANDARD: PRT: 3083 AA.
AC 089330;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Genome polypeptide [contains: N-terminal protein (P1); Helper
component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
(6K2); Genome-linked inclusion protein (VPG); Nuclear inclusion protein A (NI-A)
(NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
(EC 2.7.7.48); Coat protein (CP)].
DE Zucchini yellow mosaic virus (strain Reunion Island) (ZYMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Potyviruses.
OX NCBI_TaxID-117129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95146958; PubMed-7844540;
RA Wisler G.C., Purcell D.E., Hiebert E.;
RT *Characterization of the P1 protein and coding region of the zucchini
RT yellow mosaic virus*;
RL J. Gen. Virol. 76:37-45(1995).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
CC further restricted by preferences for the amino acids in p6 - p1'
CC that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC Glu(Ser or Gly) for the enzyme from tobacco etch virus. The
CC natural substrate is the viral polyprotein, but other proteins and
CC oligopeptides containing the appropriate consensus sequence are
CC also cleaved.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-I-Gly bond at its own C-
CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-I-Gly, in the
CC processing of the polyviral polyprotein.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L29569; AAA65558.1;
CC DR HESP: P27958; 1HEI.
CC DR MEROPS: C04.003; -.
CC DR InterPro: IPR001410; DEAD.
CC DR InterPro: IPR001650; Helicase_C.
CC DR InterPro: IPR001730; Peptidase_C4.
CC DR InterPro: IPR001456; Peptidase_C6.
CC DR InterPro: IPR001592; Poly-coat.
CC DR InterPro: IPR002540; Poly_p1.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_PTD.
CC DR InterPro: IPR007094; RNA_pol_PSVIR.
CC DR Pfam: PF00271; Helicase_C.1.
CC DR Pfam: PF00663; Peptidase_C4.1.

```

```

DR Pfam: PF00851; Peptidase_C6.1.
DR Pfam: PF00767; Poly-coat.1.
DR Pfam: PF01577; Poly_p1.1.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR PRINTS: PR00966; NIAPOTPYTASE.
DR SMART: SM00487; DEXDC.1.
DR SMART: SM00490; HELIC_C.1.
KW Hydrolyase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 307 N-TERMINAL PROTEIN.
FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.
FT CHAIN 770 1115 PROTEIN P3.
FT CHAIN 1116 1167 6 kDa PROTEIN 1.
FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1802 1854 6 kDa PROTEIN 2.
FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.
FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2288 2804 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2805 3083 COAT PROTEIN.
FT NP_BIND 1252 1259 ATP (POTENTIAL).
SQ SEQUENCE 3083 AA; 351156 MW; 55E51B455C20C537 CRC64;

Query Match          94.1%; Score 32; DB 1; Length 3083;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMLH 5
DB 2182 SYMIM 2186

RESULT 13
POLG_ZYMWV STANDARD: PRT: 3083 AA.
AC 036979;
DT 15-JUL-1999 (rel. 38, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Genome polypeptide [contains: N-terminal protein (P1); Helper
component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
(6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
(NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
(EC 2.7.7.48); Coat protein (CP)].
DE Zucchini yellow mosaic virus (strain Singapore) (ZYMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Potyviruses.
OX NCBI_TaxID-117130;
RN [1]
RP SEQUENCE FROM N.A.
RX Lee K.C., Wong S.M.;
RT Submitted (MAY-2001) to the EMBL/Genbank/DBD databases.
RT -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
RT TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
CC further restricted by preferences for the amino acids in p6 - p1'
CC that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC Glu(Ser or Gly) for the enzyme from tobacco etch virus. The
CC natural substrate is the viral polyprotein, but other proteins and
CC oligopeptides containing the appropriate consensus sequence are
CC also cleaved.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-I-Gly bond at its own C-
CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-I-Gly, in the
CC processing of the polyviral polyprotein.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE

```

POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.

-1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.

-1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.

-1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; AF014811; AAB72004.2; -

DR HSSP; P27958; 1HEI.

DR MEROPS; C04.003; -

DR MEROPS; C06.001; -

DR MEROPS; S30.001; -

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR001730; Peptidase_C4.

DR InterPro; IPR001456; Peptidase_C6.

DR InterPro; IPR001592; Poly_coat.

DR InterPro; IPR002540; Poly_P1.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVLr.

DR Pfam; PF008271; helicase_C.1.

DR Pfam; PF00863; peptidase_C4.1.

DR Pfam; PF00851; peptidase_C6.1.

DR Pfam; PF00767; Poly_P1.1.

DR Pfam; PF00577; Poly_P1.1.

DR Pfam; PF00680; RNA_dep_RNA_pol.1.

DR PRINTS; PR00966; NIAPOTYRASE.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00480; HELIC; 1.

KW Hydroxylase; transferase; 1.

KW Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase;

KW ATP-binding.

FT CHAIN 1 307 N-TERMINAL PROTEIN.

FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.

FT CHAIN 770 1115 PROTEIN P3.

FT CHAIN 1116 1167 6 kDa PROTEIN 1.

FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.

FT CHAIN 1802 1854 6 kDa PROTEIN 2.

FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.

FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.

FT CHAIN 2288 2804 NUCLEAR INCLUSION PROTEIN B.

FT CHAIN 2805 3083 COAT PROTEIN.

FT NP_BIND 1252 1259 ATP (POTENTIAL).

FT NP_BIND 3083 AA; 351028 MW; B4B6C53C1452A488 CRC64;

Query Match 94.1%; Score 32; DB 1; Length 3083;

Best Local Similarity 80.0%; Pred. No. 5.1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMLH 5

DB 2182 SYMLH 2186

RESULT 14

YESP_BACSU STANDARD; PRT; 309 AA.

AC 031519;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable ABC transporter permease protein yesp.

GN YESP.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertoro M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borries R., Bourcier L., Brans A., Braun M., Brigneau S.C., Bron S.,

RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Eutlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G.,

RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Moore D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,

RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

RA Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Wetzinger T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

subtilis".

RL Nature 390:249-256(1997).

CC -1- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM.

CC PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE SUBSTRATE ACROSS

CC THE MEMBRANE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT

CC SYSTEM PERMEASE FAMILY. MALG SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; Z99107; CAB12517.1; -

DR PIR; G69796; G69796.

DR Subtilist; BG12850; yesp.

DR InterPro; IPR000515; BPD_transp.

DR Pfam; PF00528; BPD_transp.1.

DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER.1.

DR Hypothetical protein; Transport; Transmembrane; Complete proteome.

FT TRANSMEM 29 49 POTENTIAL.

FT TRANSMEM 84 104 POTENTIAL.

FT TRANSMEM 114 134 POTENTIAL.

FT TRANSMEM 167 187 POTENTIAL.

FT TRANSMEM 217 237 POTENTIAL.

FT TRANSMEM 275 295 POTENTIAL.

FT SEQUENCE 309 AA; 34614 MW; 971DE02DA4F9153B CRC64;

Query Match 91.2%; Score 31; DB 1; Length 309;

Best Local Similarity 80.0%; Pred. No. 83;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMLH 5

DB 2182 SYMLH 2186

Db 299 SYWVH 303

RESULT 15

TAM2_HUMAN STANDARD; PRT; 370 AA.

ID TAM2_HUMAN

AC 015035;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Translocation associated membrane protein 2.

GN TRAM2 OR KIAA0057.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

ON [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=96051398; PubMed=7584044;

RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S., Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;

RT "Prediction of the coding sequences of unidentified human genes. II. The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by analysis of cDNA clones from human cell line KG-1.";

RT DNA Res. 1:223-229(1994).

RM [2]

RP SEQUENCE FROM N.A.

RA Tracey A.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=2388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Dadin T.B., Toshilyuk S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez R., Bouffard G.G., Whitting R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE LASSI FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D31762; BAA06540.1; -

DR EMBL; AL049611; CAB71119.1; -

DR EMBL; BC028121; AAH28121.1; -

DR Genew; HGNC:16855; TRAM2.

DR Interpro; IPR005547; LAG1.

DR Interpro; IPR006634; TIG.

DR Pfam; PF03798; LAG1; 1.

DR SMART; SM00724; TIG; 1.

KW Transmembrane; Glycoprotein; Translocation.

FT TRANSMEM 23 43 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 120 140 POTENTIAL.

FT TRANSMEM 160 180 POTENTIAL.

FT TRANSMEM 199 219 POTENTIAL.

FT TRANSMEM 251 271 POTENTIAL.

FT TRANSMEM 288 308 POTENTIAL.

FT CARBOHYD 55 55

SQ SEQUENCE 370 AA; 43327 MW; 9B5183F1A3D45366 CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

Query Match 91.2%; Score 31; DB 1; Length 370;

Best Local Similarity 80.0%; Pred. No. 99;

Matches 4; Conservative 1; Mismatches 0; Gaps 0;

QY 1 SYWVH 5

Db 169 AYWVH 173

Search completed: October 7, 2003, 19:15:01

Job time : 2.01724 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 4.98276 Seconds
(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-4_COPY_31_35

Perfect score: 34
Sequence: 1 STWLH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_rv1rus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	90	16	09c142
2	34	100.0	105	10	08w0f7
3	34	100.0	196	10	04g198
4	34	100.0	224	11	09d561
5	34	100.0	250	10	09xev8
6	34	100.0	302	11	09d592
7	34	100.0	317	10	09sae0
8	34	100.0	430	5	026424
9	34	100.0	482	16	08y402
10	34	100.0	494	16	092es4
11	34	100.0	537	16	08x070
12	34	100.0	566	5	094013
13	34	100.0	675	17	08rnp4
14	34	100.0	4924	3	09p771
15	32	94.1	114	11	09j181
16	32	94.1	117	11	09z1c6

17	32	94.1	137	11	0924r6	0924r6 mus musculus
18	32	94.1	139	11	0924r5	0924r5 mus musculus
19	32	94.1	140	11	0924p8	0924p8 mus musculus
20	32	94.1	140	11	0924r2	0924r2 mus musculus
21	32	94.1	141	11	0924o4	0924o4 mus musculus
22	32	94.1	142	11	0924o1	0924o1 mus musculus
23	32	94.1	143	11	091va2	091va2 mus musculus
24	32	94.1	143	11	0924o5	0924o5 mus musculus
25	32	94.1	143	11	091v67	091v67 mus musculus
26	32	94.1	143	11	0924r7	0924r7 mus musculus
27	32	94.1	143	11	0924p6	0924p6 mus musculus
28	32	94.1	143	11	0924r0	0924r0 mus musculus
29	32	94.1	144	11	0924p5	0924p5 mus musculus
30	32	94.1	145	11	0924o6	0924o6 mus musculus
31	32	94.1	145	11	0924o9	0924o9 mus musculus
32	32	94.1	145	11	0924r3	0924r3 mus musculus
33	32	94.1	145	11	0924o7	0924o7 mus musculus
34	32	94.1	145	11	0924p7	0924p7 mus musculus
35	32	94.1	145	11	0924r1	0924r1 mus musculus
36	32	94.1	145	11	0924r4	0924r4 mus musculus
37	32	94.1	146	11	0924r8	0924r8 mus musculus
38	32	94.1	146	11	0924o3	0924o3 mus musculus
39	32	94.1	189	5	046120	046120 fasciola he
40	32	94.1	278	11	0921k1	0921k1 mus musculus
41	32	94.1	371	5	0903p5	0903p5 caenorhabdi
42	32	94.1	373	5	09xxk7	09xxk7 caenorhabdi
43	32	94.1	434	8	09gg28	09gg28 naies minor
44	32	94.1	466	12	08jxj9	08jxj9 tomato spot
45	32	94.1	466	12	08jxk0	08jxk0 tomato spot

ALIGNMENTS

RESULT 1

09c142 ID 09c142 PRELIMINARY; PRT; 90 AA.
AC 09c142;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Unknown protein.
GN YRFG OR L0524.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IL1403;
RX MEDLINE-21235186; PubMed-11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarre K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001).
DR EMBL; AF006287; AAK04622.1; -.
KW Complete Proteome.
SQ SEQUENCE 90 AA; 10209 MW; 1D4A61B0F3FF84P9 CRC64;

Query Match Score 34; DB 16; Length 90;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STWLH 5
Db 17 STWLH 21

RESULT 2

08w0f7 ID 08w0f7 PRELIMINARY; PRT; 105 AA.
AC 08w0f7;
DT 01-MAR-2002 (Tremblrel. 20, Created)

```

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE P0529E05.22 protein.
GN P0529E05.22.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriacridaceae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Saeki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(Gs3) genomic DNA, chromosome 1, PAC
RT clone:P0529E05."
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP003279; BAB84406.1; -.
DR Gramene; O8W0E7; -.
SQ SEQUENCE 105 AA; 12386 MW; 2BA42D3A6D61620F CRC64;

Query Match 100.0%; Score 34; DB 10; Length 105;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 50 SYWLH 54

RESULT 3
ID 049198 PRELIMINARY; PRT; 196 AA.
AC 049198;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DR 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DR 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hairy root 4.
GN HR4.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hairy root tumor;
RA Nakamura C., Nishimura N., Mishima M., Ueda T., Onaka T.,
RA Tavelanov S., Takumi S., Mori N.;
RT "Molecular cloning and characterization of cDNAs preferentially
RT expressed during the development of hairy roots and normal roots in
RT tobacco (Nicotiana tabacum L.)."
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF043554; AAC02087.1; -.
DR HSSP; P24337; IHRP.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryr_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 196 AA; 19904 MW; 7225COBBA1BEFB5 CRC64;

Query Match 100.0%; Score 34; DB 10; Length 196;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 153 SYWLH 157

RESULT 4
ID 09D561 PRELIMINARY; PRT; 224 AA.
AC 09D561;
DR 01-JUN-2001 (TREMBlrel. 17, Created)

```

```

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 493051M1IRIK protein.
GN 493051M1IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Testis;
RX MEDLINE=11085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachii J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochika H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.W., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Notone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK015761; BAB29963.1; -.
DR MGD; MGI:1922305; 493051M1IRIK.
DR InterPro; IPR006214; UPF0005.
DR Pfam; PF01027; UPF0005; 1.
SQ SEQUENCE 224 AA; 25646 MW; 6367E0BBB8BD93C CRC64;

Query Match 100.0%; Score 34; DB 11; Length 224;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 160 SYWLH 164

RESULT 5
ID 09XEV8 PRELIMINARY; PRT; 290 AA.
AC 09XEV8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DR 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DR 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE CCAT-binding transcription factor subunit B.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC Mundree S.G., Singh N.P., Singh N.K.;
RT "Cloning of two novel tobacco cDNA clones by functional sufficiency
RT for salt tolerance in Escherichia coli."
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF120092; AAD28439.1; -.
DR InterPro; IPR001289; TF_CBBF.
DR Pfam; PF02045; CBBF_NFYA; 1.
DR PRINTS; PR00616; CCAATSUBUNT.
DR ProDom; PD003860; TF_CBBF; 1.
DR SMART; SM00521; CBBF_1.
SQ SEQUENCE 290 AA; 33949 MW; 5F275974A10AF931 CRC64;

```

Query Match 100.0%; Score 34; DB 10; Length 290;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
 |||||
 DB 183 SYWLH 187

RESULT 6

Q9D592 PRELIMINARY; PRT; 302 AA.
 ID Q9D592
 AC Q9D592;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE 4930500J03RLK protein.
 GN 4930500J03RLK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh J., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Alzawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gstaalich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmberg L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK015666; BAB29920.1; -
 DR MGI; MGI:1922260; 4930500J03RLK.
 DR InterPro; IPR006214; UPE0005.
 DR Pfam; PF01027; UPE0005; 1.
 SQ SEQUENCE 302 AA; 34184 MW; CFE5E4B9081554C CRC64;

Query Match 100.0%; Score 34; DB 11; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
 |||||
 DB 238 SYWLH 242

RESULT 7

Q9SAE0 PRELIMINARY; PRT; 317 AA.
 ID Q9SAE0
 AC Q9SAE0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F3F19.7.
 GN F3F19.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lee J.M., Lanz C.,
 RA Liu S., Li J., Kremetska I., Lucos J., Ngan I., Gonzalez A.,
 RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
 RA Hansen N., Huizler L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F3F19 sequence.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Theologis A.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007357; AAD31058.1; -
 SQ SEQUENCE 317 AA; 35809 MW; 17972E206698CCD8 CRC64;

Query Match 100.0%; Score 34; DB 10; Length 317;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
 |||||
 DB 289 SYWLH 293

RESULT 8

Q26424 PRELIMINARY; PRT; 430 AA.
 ID Q26424
 AC Q26424;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE C-myc proto-oncogene homolog protein (Fragment).
 OS Crassostrea virginica (Eastern oyster).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;
 OC Ostreidae; Ostreidae; Crassostrea.
 OX NCBI_TaxID=6565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95291374; PubMed=7773336;
 RA Marsh A.G., Chen T.T.;
 RT "A divergent cDNA homologue of the c-myc proto-oncogene in the eastern
 oyster Crassostrea virginica: Implications for Myc evolution.";
 RL Mol. Mar. Biol. Biotechnol. 4:185-192(1995).
 DR EMBL; S77334; AAB34577.1; -
 DR InterPro; IPR003582; SHKT.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01549; SHKT; 2.
 DR Pfam; PF00093; VWC; 1.
 DR SMART; SM00254; SHKT; 2.
 DR SMART; SM00214; VWC; 1.
 FT NON TER 1
 SQ SEQUENCE 430 AA; 47896 MW; DPAFCF3A9B32E5DA3 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWLH 5
 |||||
 DB 240 SYWLH 244

RESULT 9

O8YA02 PRELIMINARY; PRT; 482 AA.
 ID O8YA02
 AC O8YA02;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE Hypothetical protein lmo0365.
 GN LMO0365.
 OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillales; *Listeriaceae*; *Listeria*.
 NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BGD-e / Serovar 1/2a;
 RX MEDLINE-21537279; PubMed-11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
 Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurgey O.,
 Entlan K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordtek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL: AL591975; CAC98444.1; -
 DR ListList; LMO0365; -
 DR InterPro: IPR003438; ABC_transporter.
 DR InterPro: IPR004923; FTR1.
 DR Pfam: PF03239; FTR1; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 482 AA; 53669 MW; D8ADD8062990EDF CRC64;
 Query Match 100.0%; Score 34; DB 16; Length 482;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SWMLH 5
 Db 314 SWMLH 318

RESULT 10
 092ES4 PRELIMINARY; PRT; 494 AA.
 AC 092ES4;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
 DE Hypothetical protein lln0384.
 GN Lln0384.
 OS *Listeria innocua*.
 OC Bacteria; Firmicutes; Bacillales; *Listeriaceae*; *Listeria*.
 NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CUP 11262 / Serovar 6a;
 RX PubMed-11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
 Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurgey O.,
 Entlan K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordtek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL: AL596164; CAC95617.1; -
 DR ListList; Lln0384; -

DR InterPro: IPR004923; FTR1.
 DR Pfam: PF03239; FTR1; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 494 AA; 55050 MW; 882E6F1BF688E32 CRC64;
 Query Match 100.0%; Score 34; DB 16; Length 494;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SWMLH 5
 Db 314 SWMLH 318

RESULT 11
 08X070 PRELIMINARY; PRT; 537 AA.
 ID 08X070
 AC 08X070;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)
 DE Probable transmembrane protein.
 GN RSP1416 OR RS03125.
 OS *Ralstonia solanacearum* (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; *Ralstonia*.
 NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CMI1000;
 RX MEDLINE-21681879; PubMed-11823852;
 RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brotier P., Camus J.-C., Catolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*."
 RL Nature 415:497-502(2002).
 DR EMBL: AL646084; CAD18567.1; -
 KM Plasmid; Complete proteome.
 SQ SEQUENCE 537 AA; 57305 MW; FFF20EDA34ABE3D CRC64;
 Query Match 100.0%; Score 34; DB 16; Length 537;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SWMLH 5
 Db 41 SWMLH 45

RESULT 12
 094013 PRELIMINARY; PRT; 566 AA.
 AC 094013;
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE T08G11.4 protein.
 GN T08G11.4.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Pelodierinae; *Caenorhabditis*.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dobson R.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;


```

RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
   investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z80220; CAB02307.2; -.
DR WormPep: T08611.4; CE23958.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR003402; Unk_Met10.
DR Pfam: PF02475; Met_10; 1.
SQ SEQUENCE 566 AA; 65446 MW; E59BB1E0A5BC682A CRC64;

Query Match      100.0%; Score 34; DB 5; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 133 SYWLH 137

RESULT 13
O8TNP4 PRELIMINARY; PRT; 675 AA.
AC O8TNP4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein MA2239.
GN MA2239.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
   and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL: AE010910; AAM05634.1; -.
DR InterPro: IPR001237; Postsynaptic.
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR; 8.
DR ProDom: PD012428; Postsynaptic; 1.
DR SMART: SM00028; TPR; 8.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 675 AA; 77117 MW; 551931ADC3EC8680 CRC64;

Query Match      100.0%; Score 34; DB 17; Length 675;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 230 SYWLH 234

RESULT 14
O9P7T1 PRELIMINARY; PRT; 4924 AA.
AC O9P7T1;

```

```

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative peptide synthetase, with 3 Phosphopantetheine attachment
   sites.
GN SPAC2363.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL138854; CAB72227.1; -.
DR HSSP: P14687; IAMP.
DR GeneDB_Spombe: SPAC2363.02c; -.
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR000873; AMP_bind.
DR InterPro: IPR001242; Condensatn.
DR InterPro: IPR006163; PP_bind.
DR Pfam: PF00501; AMP-binding; 4.
DR Pfam: PF00668; Condensation; 4.
DR Pfam: PF00550; PP-binding; 4.
DR PROSITE: PS50075; ACP_DOMAIN; 6.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE: PS00455; AMP_BINDING; 2.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.
SQ SEQUENCE 4924 AA; 559835 MW; 985D30FED3406E2C CRC64;

Query Match      100.0%; Score 34; DB 3; Length 4924;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWLH 5
DB 1794 SYWLH 1798

RESULT 15
O9JLB1 PRELIMINARY; PRT; 114 AA.
AC O9JLB1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Anti-myosin Immunoglobulin heavy chain variable region
   (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C;
RX MEDLINE=20448942; PubMed=10992488;
RA Malik S., Liao L., Cunningham M.W., Diamond B.;
RT "T-cell-dependent antibody response to the dominant epitope of
   streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
   with cardiac myosin.";
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL: AF206025; AAF6323.1; -.
DR HSSP: P01810; 2FSD.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1

```

FT NON_TER 114 114
SQ SEQUENCE 114 AA: 12829 MM: 40485FDE6BA56F8 CRC64;
Query Match 94.1%; Score 32; DB 11; Length 114;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWLH 5
111:1
DB 23 SYWMH 27

Search completed: October 7, 2003, 19:20:01
Job time : 5.98276 secs


```

XX 12-AUG-1994; 94US-0289576.
XX (IMMU-) IMMUNOMEDICS INC.
XX Hansen H, Leung S;
XX WPI; 1996-139454/14.
XX N-PSDB; AAT15802.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
XX for the therapy and diagnosis of B-cell lymphoma(s) and
XX leukemia(s).
XX
XX Claim 5; Page 36-37; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
XX antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
XX recombinantly linked to the framework sequences of human VK and VH
XX regions, respectively, to give humanised LL2 VK (AAR92217) and VH
XX (AAR92218). These were subsequently linked, respectively, to human
XX kappa and IgG1 constant regions. A humanised Mab was obtd. that
XX retained the B-lymphoma and leukaemia cell targeting and
XX internalisation characteristics of the parental LL2 Mab, and which
XX exhibited a lowered HAMA reaction. It can be linked to e.g. a
XX cytostatic agent for therapeutic appln.
XX
XX Sequence 116 AA;
XX
XX Query Match 100.0%; Score 99; DB 17; Length 116;
XX Best Local Similarity 100.0%; Pred. No. 6e-08;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 YINPRNDYTEYXNONFKD 17
XX |||||||||||||||
XX 50 YINPRNDYTEYXNONFKD 66
XX
XX RESULT 2
XX AAR92218
XX ID AAR92218 standard; Protein; 116 AA.
XX AC AAR92218;
XX AC AAR92218;
XX DT 28-MAY-1996 (first entry)
XX DE Humanised LL2 Mab VH region.
XX DE
XX KM Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
XX KM leukaemia; therapy; diagnosis; complementarity determining region;
XX KM CDR; antibody engineering.
XX OS Synthetic.
XX OS
XX FH Key Location/Qualifiers
XX FH Region 31..35
XX FT /label= CDR1
XX FT Region 50..66
XX FT /label= CDR2
XX FT Region 99..105
XX FT /label= CDR3
XX FT
XX PN WO9604925-A1.
XX PN
XX PD 22-FEB-1996.
XX PD
XX PF 11-AUG-1995; 95WO-US09641.
XX PF
XX PR 12-AUG-1994; 94US-0289576.
XX PR
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PA
XX PT Hansen H, Leung S;
XX PT

```

```

XX WPI; 1996-139454/14.
XX N-PSDB; AAT15804.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
XX for the therapy and diagnosis of B-cell lymphoma(s) and
XX leukemia(s).
XX
XX Claim 5; Page 39; 70pp; English.
XX
XX The complementarity determining regions (CDRs) of mouse monoclonal
XX antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
XX recombinantly linked to the framework sequences of human VK and VH
XX regions, respectively, to give humanised LL2 VK (AAR92217) and VH
XX (AAR92218). These were subsequently linked, respectively, to human
XX kappa and IgG1 constant regions. A humanised Mab was obtd. that
XX retained the B-lymphoma and leukaemia cell targeting and
XX internalisation characteristics of the parental LL2 Mab, and which
XX exhibited a lowered HAMA reaction. It can be linked to e.g. a
XX cytostatic agent for therapeutic appln.
XX
XX Sequence 116 AA;
XX
XX Query Match 100.0%; Score 99; DB 17; Length 116;
XX Best Local Similarity 100.0%; Pred. No. 6e-08;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 YINPRNDYTEYXNONFKD 17
XX |||||||||||||||
XX 50 YINPRNDYTEYXNONFKD 66
XX
XX RESULT 3
XX AAR92219
XX ID AAR92219 standard; Protein; 116 AA.
XX AC AAR92219;
XX AC AAR92219;
XX DT 28-MAY-1996 (first entry)
XX DE Humanised LL2 Mab VH region (version hLL2-1).
XX DE
XX KM Humanised antibody; monoclonal antibody; Mab; LL2; B-cell lymphoma;
XX KM leukaemia; therapy; diagnosis; complementarity determining region;
XX KM CDR; antibody engineering.
XX OS Synthetic.
XX OS
XX FH Key Location/Qualifiers
XX FH Region 31..35
XX FT /label= CDR1
XX FT Region 50..66
XX FT /label= CDR2
XX FT Region 99..105
XX FT /label= CDR3
XX FT
XX PN WO9604925-A1.
XX PN
XX PD 22-FEB-1996.
XX PD
XX PF 11-AUG-1995; 95WO-US09641.
XX PF
XX PR 12-AUG-1994; 94US-0289576.
XX PR
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PA
XX PT Hansen H, Leung S;
XX PT
XX WPI; 1996-139454/14.
XX WPI; 1996-139454/14.
XX
XX Chimeric and humanised LL2 antibodies - used to produce conjugates
XX for the therapy and diagnosis of B-cell lymphoma(s) and
XX leukemia(s).
XX

```

```

XX Example 1: Page 40; 70pp; English.
PS
XX The complementarity determining regions (CDRs) of mouse monoclonal
CC antibody (MAb) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC recombinantly linked to the framework sequences of human VK and VH
CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC (AAR92218). In an alternative version, a glutamine was introduced
CC at position 5 of the humanised VH (AAR92219) to include a PstI site
CC useful for subcloning. The humanised VK and VH were subsequently
CC linked, respectively, to human kappa and IgG1 constant regions. A
CC humanised MAb was obtained that retained the B-lymphoma and leukaemia
CC cell targeting and internalisation characteristics of the parental
CC LL2 MAb, and which exhibited a lowered HAMA reaction. It can be
CC linked to a cytostatic agent for therapeutic appln.
CC
XX
SQ Sequence 116 AA:

Query Match 100.0%; Score 99; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
   |||||||||||||||
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 4
AAW27696
ID AAW27696 standard; Protein; 116 AA.
AC
XX AAW27696;
AC
XX
DT 14-APR-1998 (first entry)
XX
XX Variable heavy chain of MAb LL2.
DE
XX
XX Variable heavy chain; B cell; monoclonal antibody; MAb; LL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine;
KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
XX Mus sp.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Region 31..35
FT /note= "complementarity determining region 1"
FT 50..66
FT /note= "complementarity determining region 2"
FT 99..105
FT Region /note= "complementarity determining region 3"
XX
XX WO9734632-A1.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04196.
XX
XX 20-MAR-1996; 96US-0013709.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen H, Leung S, Qu Z;
XX
XX WPI: 1997-479995/44.
XX
XX N-PSDB; AAT88129.
XX
XX Monoclonal antibody engineered to contain glycosylation site - in
XX non-Fc constant heavy or light chain region, useful to diagnose or
XX treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
XX Example 3; Fig 4B; 88pp; English.
PS

```

```

XX
XX The present sequence is the variable heavy chain of the
CC B cell specific monoclonal antibody (MAb) LL2, which contains an
CC engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly
CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
CC murine MAb. The MAb can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia. The glycosylation site allows a label or
CC therapeutic agent of increased size to be conjugated to the
CC carbohydrate moiety, without affecting the MAb's binding affinity
CC or specificity.
XX
XX
SQ Sequence 116 AA:

Query Match 100.0%; Score 99; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
   |||||||||||||||
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 5
AAW27698
ID AAW27698 standard; Protein; 116 AA.
AC
XX AAW27698;
AC
XX
DT 14-APR-1998 (first entry)
XX
XX Variable heavy chain of MAb hLL2.
DE
XX
XX Variable heavy chain; B cell; monoclonal antibody; MAb; hLL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine; humanised;
KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
XX Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 31..35
FT /note= "complementarity determining region 1"
FT 50..66
FT /note= "complementarity determining region 2"
FT 99..105
FT Region /note= "complementarity determining region 3"
XX
XX WO9734632-A1.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04196.
XX
XX 20-MAR-1996; 96US-0013709.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen H, Leung S, Qu Z;
XX
XX WPI: 1997-479995/44.
XX
XX N-PSDB; AAT88131.
XX
XX Monoclonal antibody engineered to contain glycosylation site - in
XX non-Fc constant heavy or light chain region, useful to diagnose or
XX treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
XX Example 3; Fig 5B; 88pp; English.
XX
XX The present sequence is the variable heavy chain of the
XX B cell specific monoclonal antibody (MAb) hLL2. hLL2 is a highly
XX specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
XX
XX

```

CC humanised murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukaemia.

XX Sequence 116 AA;

Query Match 100.0%; Score 99; DB 18; Length 116;

Best Local Similarity 100.0%; Pred. No. 6e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNONFKD 17

DB 50 YINPRNDYTEYNONFKD 66

RESULT 6

AAAP3079

ID AAP93079 standard; peptide; 159 AA.

XX AAP93079;

AC 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 15-MAR-1990 (first entry)

XX Heavy chain of monoclonal antibody 6A4.

DE Monoclonal antibody 6A4; heavy chain; Pseudomonas aeruginosa; OMP-1.

XX Unidentified.

OS EP38395-A.

XX 25-OCT-1989.

PD 12-APR-1989; 89EP-0106463.

XX 19-APR-1988; 88DE-3613023.

PR (BEHW) BEHRINGER AG.

XX Domdey H, Marget M, Vonspecht B;

PI WPI: 1989-310861/43.

DR N-PSDB; AAN91645.

XX Monoclonal antibody to Pseudomonas aeruginosa - and DNA coding for

PT variable antibody regions.

XX Claim 1; page 6; 7pp; german.

PS The peptide is encoded by the heavy chain of monoclonal antibody 6A4.

XX 6A4 reacts with the OMP-1 protein of all 19 known serotypes of

CC P.aeruginosa. It is used for therapy and diagnosis of infection, and as

CC a carrier for drugs. The antibody is IgG2a subclass.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

CC Sequence 159 AA;

QY Query Match 81.8%; Score 81; DB 10; Length 159;

Best Local Similarity 82.4%; Pred. No. 5.9e-05;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNONFKD 17

DB 69 YINPRNDYTEYNONFKD 85

RESULT 7

AAAR21278

ID AAR21278 standard; Protein; 114 AA.

XX AAR21278;

XX 21-MAY-1992 (first entry)

XX Murine VH group 1 chain O specific for phox.

DE Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat;

XX plus; gp3; binding; adsorption; gene VIII; diverse repertoire;

XX specific binding pairs; replicable genetic display package.

XX Synthetic.

OS Key Location/Qualifiers

XX Binding-site 31..35

FT /label- CDR1

FT Binding-site 50..66

FT /label- CDR2

FT Binding-site 99..103

FT /label- CDR3

FT /note=" D/N-X-G-X-X motif "

XX W09201047-A.

XX 23-JAN-1992.

XX 10-JUL-1991; 91WO-GB01134.

XX 15-MAY-1991; 91GB-0010549.

XX 10-JUL-1990; 90GB-0015198.

XX 19-OCT-1990; 90GB-0022845.

XX 12-NOV-1990; 90GB-0024503.

XX 06-MAR-1991; 91GB-0004744.

XX (CAMP-) CAMBRIDGE ANTIBODY.

XX (MEDI-) MED RES COUNCIL.

XX McCafferty J, Pope AR, Johnson KS, Hoogenboom HRJ, Griffiths AD;

XX Jackson RH, Holliger KP, Marks JD, Clackson TP, Chiswell DJ;

XX Winter GP, Bonnett TP;

XX WPI: 1992-056862/07.

XX Producing members of specific binding pairs - by expression in

PT recombinant host cells with a secreting replicable genetic

PT display package.

XX Example 22; Fig 24; 209pp; English.

XX The VH sequence is one of 22 (AAR21264-85) expressed from a single

CC chain Fv library. The library produces a diverse repertoire of

CC antibody fragments specific for 2-phenyl-5-oxazolone (phox). It was

CC prep'd. using cDNA generated from mRNA from mice immunised with phox

CC coupled to chicked serum albumin. The VH and VL kappa sequences

CC were separately amplified by PCR (AAG23474-84) and ligated into

CC fcdm2 (AAG23463) for expression on the phage surface as fusions with

CC gene III. The resulting library of clones was diverse; 23 hapten

CC binding clones were sequenced revealing 8 VH genes (A-H; AAR21264-71)

CC in a variety of pairings with 7 VK genes (a-g; AAR21286-92). Most

CC clones were VK-d combinations so a further hierarchical library was

CC prep'd. by "crossing" VK-d with the VH repertoire. The resulting

CC library was screened for hapten binding and 24 clones sequenced. 13

CC new partners (AAR21272-85) for VK-d were identified. Nearly all the

CC VH genes belonged to gp 1, with only one, "E", being of gp 2 (VHox1).

CC Of the 24 hierarchical clones, only one was of type "O". The Kd of

CC VH-B/VK-d for phox-GABA was 10 nM, one of the highest values found.

CC This suggests that phage bearing scfv fragments having weak affinity

CC can be selected with antigen, probably due to the avidity of the

CC multiple antibody heads on the phage. The different combinations

CC could also be isolated on a basis of antigen affinity.

CC See also AAR21260-307, 309-311; AAR22450, 565-581.

XX Sequence 114 AA;

Query Match 80.8%; Score 80; DB 13; Length 114;
 Best Local Similarity 82.4%; Pred. No. 5.9e-05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQRFKD 17
 ||||| ||||| |||
 Db 50 YINPRTGYTEYNQRFKD 66

RESULT 8

AAW76127
 ID AAW76127 standard; Protein; 138 AA.

AC AAW76127;
 XX
 XX 20-NOV-1998 (first entry)
 XX

DE Murine ICR-1.1 V-H region PCR product protein.

XX Intercellular adhesion molecule; ICAM-R; modulator; 14.3.3 family;
 KW HSI-beta; tubulin; inhibitor; stimulator; effector; immune response;
 KW inflammation; disorder; T cell activation; macrophage; Crohn's disease;
 KW adult respiratory distress syndrome; stroke; multiple sclerosis; asthma;
 KW rheumatoid arthritis; tumour growth; human immune deficiency virus;
 KW infection; diabetes; graft vs. host disease; passive immunisation.

XX Mus sp.

XX Key Location/Qualifiers
 FH Misc-difference 2 /Label= unknown
 FT Misc-difference 6 /Label= unknown

FT Misc-difference 7 /Label= unknown
 FT Misc-difference 7 /Label= unknown

FT Misc-difference 8 /Label= unknown
 FT

XX US5773218-A.

XX 30-JUN-1998.

XX 07-JUN-1995; 95US-0482882.

XX 05-AUG-1994; 94US-0286754.

XX 27-JAN-1992; 92US-0827689.

XX 26-MAY-1992; 92US-0889724.

XX 05-JUN-1992; 92US-0894061.

XX 22-JAN-1993; 93US-0009266.

XX 26-JAN-1993; 93WO-US00787.

XX 05-AUG-1993; 93US-0102852.

XX 07-JUN-1995; 95US-0482882.

XX (ICOS-) ICOS CORP.

XX Gallatin WM, Vazeux R;

XX WPI: 1998-386989/33.

XX N-PSDB: AAV56413.

XX Identifying compounds that modulate interaction of intercellular

XX adhesion molecule R - with ligands HSI-beta and tubulin using

XX two-hybrid assay, useful for treating inflammation, T cell

XX activation etc.

XX Example 13: Column 125-128; 108bp; English.

CC antibody optionally coupled to toxin or radionuclide, or an ICAM-R
 CC peptide, can block, inhibit or stimulate ligand/receptor interactions
 CC involving ICAM-R, particularly its effector functions involved in
 CC (non)specific immune responses. ICAM-R related agents may be used to
 CC treat or monitor inflammation, disorders involving T cell activation or
 CC macrophages, e.g. adult respiratory distress syndrome, stroke, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis, asthma, tumour growth,
 CC human immune deficiency virus infection, diabetes, graft vs. host disease
 CC and many others. Antibodies may also be used for passive immunisation,
 CC for purifying, detecting or quantifying ICAM-R and for identifying
 CC ICAM-R expressing cells.

XX Sequence 138 AA;

Query Match 78.8%; Score 78; DB 19; Length 138;
 Best Local Similarity 76.5%; Pred. No. 0.00015;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQRFKD 17
 ||||| ||||| |||
 Db 69 YINPNTGYTEYNQRFKD 85

RESULT 9
 AAW71257
 ID AAW71257 standard; Protein; 138 AA.

XX AAW71257;

XX 25-MAR-2003 (updated)
 DT 18-NOV-1998 (first entry)
 XX

DE Murine antibody ICR-1.1 heavy chain amino acid sequence.

XX Human; ICAM-R; intercellular adhesion molecule; adhesion; treatment;
 KW inflammatory condition; asthma; tumour growth; metastasis;
 KW viral infection; antibody ICR-1.1.

XX Mus sp.

XX Key Location/Qualifiers
 FH Misc-difference 2 /note= "encoded by GRA"

FT Misc-difference 6 /note= "encoded by RTC"

FT Misc-difference 7 /note= "encoded by WTB"

FT Misc-difference 8 /note= "encoded by HTC"

FT

XX US5811517-A.

XX 22-SEP-1998.

XX 07-JUN-1995; 95US-0483389.

XX 05-AUG-1994; 94US-0286754.

XX 26-JAN-1993; 93WO-US00787.

XX 27-JAN-1992; 92US-0827689.

XX 05-JUN-1992; 92US-0889724.

XX 26-MAY-1992; 92US-0894061.

XX 22-JAN-1993; 93US-0009266.

XX 05-AUG-1993; 93US-0102852.

XX (ICOS-) ICOS CORP.

XX Gallatin WM, Vazeux R;

XX WPI: 1998-530940/45.

XX N-PSDB: AAV54863.

XX DNA encoding mutant ICAM-R poly:peptide(s) - useful for diagnosis
 PT and treatment of cell adhesion based disease conditions e.g.

PT Inflammation or asthma
 PS Example 13; Columns 125-126; 11bp; English.
 XX
 CC The present sequence represents the heavy chain of murine antibody
 CC ICR-1.1. This antibody is specific for ICAM-R (intercellular adhesion
 CC molecule-R). ICAMs are polypeptides that are expressed on blood vessel
 CC endothelial cell surfaces and are involved in the adhesion events in
 CC various conditions. ICAM-R variants (see AAW71264-69) can be used to
 CC treat or monitor inflammatory conditions involving specific or
 CC non-specific immune responses, asthma, tumour growth and/or metastasis
 CC and viral infections. The ICAM variants are produced recombinantly, from
 CC expression libraries of mutated sequences, and the ones that are
 CC claimed are the ones that have been found to be especially involved in
 CC adhesion events. They can also be used to raise antibodies, also for
 CC use as therapeutic or diagnostic agents.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 SQ Sequence 138 AA;
 Query Match 78.88; Score 78; DB 19; Length 138;
 Best Local Similarity 76.58; Pred. No. 0.00015;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 YINPRNDYTEXNQFKD 17
 |||| ||||| |:
 Db 69 YINPNTDYTEXNQRFQD 85
 RESULT 10
 AAW00783
 ID AAW00783 standard; Protein: 138 AA.
 XX
 AC AAW00783;
 XX
 DT 14-MAY-1999 (first entry)
 XX
 DE Antibody against ICAM-R.
 XX
 KW ICAM; immunoglobulin-like loop; intercellular adhesion molecule receptor;
 KW alpha d/CD18; antibody; immunisation; inflammatory response; asthma;
 KW tumour growth; viral infection; therapy.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..138
 FT /note= "Xaa- unspecified amino acid"
 XX
 PN US5880268-A.
 XX
 PD 09-MAR-1999.
 XX
 PF 07-JUN-1995; 95US-0483932.
 XX
 PR 05-AUG-1994; 94US-0286754.
 PR 27-JAN-1992; 92US-0827689.
 PR 26-MAY-1992; 92US-0889724.
 PR 05-JUN-1992; 92US-0894061.
 PR 22-JAN-1993; 93US-0009266.
 PR 26-JAN-1993; 93WO-US00787.
 PR 05-AUG-1993; 93US-0102852.
 PR 07-JUN-1995; 95US-0483932.
 XX
 XX (ICOS-) ICOS CORP.
 XX
 PI Gallatin WM, Vazeux R;
 XX
 DR WPI; 1999-204041/17.
 DR N-PSDB; AAX21879.
 XX
 PT New intercellular adhesion molecule receptor (ICAM-R) specific
 PT antibodies - useful for modulating ligand/receptor binding and

PT biological activities involving ICAM-R, especially those of the
 PT specific and non-specific immune systems
 XX
 PS Example 13; Column 125-126; 108bp; English.
 XX
 CC This sequence represents an antibody specific for ICAM-R.
 CC The invention relates to antibodies (Ab) which bind specifically
 CC to the intercellular adhesion molecule receptor (ICAM-R), inhibiting the
 CC interaction between ICAM-R and alpha d/CD18. Abs with specific ICAM-R
 CC binding are useful in compositions for immunisation, and for purifying
 CC ICAM-R polypeptides and identifying cells expressing ICAM-R on their cell
 CC surface, modulating ligand/receptor binding and biological activities
 CC involving ICAM-R, especially inflammatory responses of the specific
 CC immune system, the non-specific immune system, monitoring and treating
 CC asthma, tumour growth, and/or metastasis, and viral infection (e.g. HIV
 CC infection). In particular diseases involving an essential T cell
 CC activation (e.g. asthma, psoriasis, diabetes, graft vs. host disease,
 CC tissue transplant rejection, and multiple sclerosis) may be treated with
 CC anti-ICAM-R antibodies. The Abs specifically bind to and identify ICAM-R
 CC and disrupt ICAM-R to cell adhesion molecule, especially alpha d/CD18
 CC binding.
 CC
 SQ Sequence 138 AA;
 Query Match 78.88; Score 78; DB 20; Length 138;
 Best Local Similarity 76.58; Pred. No. 0.00015;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 YINPRNDYTEXNQFKD 17
 |||| ||||| |:
 Db 69 YINPNTDYTEXNQRFQD 85
 RESULT 11
 AAW81451
 ID AAW81451 standard; Protein: 138 AA.
 XX
 AC AAW81451;
 XX
 DT 17-FEB-1999 (first entry)
 XX
 DE Murine antibody ICR-1.1 Vh region.
 XX
 KW Intercellular adhesion molecule polypeptide; ICAM-R; humanised; ICR-1.1;
 KW ICR-8.1; monoclonal antibody; therapeutic; inflammatory; asthma; tumour;
 KW graft-versus-host disease; viral infection; toxin; radionuclide;
 KW neovascularisation site; murine.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2
 FT /label= unknown
 FT /note= "encoded by GRA"
 FT
 FT Misc-difference 6
 FT /label= unknown
 FT /note= "encoded by RTC"
 FT
 FT Misc-difference 7
 FT /label= unknown
 FT /note= "encoded by WTB"
 FT
 FT Misc-difference 8
 FT /label= unknown
 FT /note= "encoded by HTC"
 FT
 FT Misc-difference 32
 FT /note= "encoded by AAA"
 XX
 PN US5837822-A.
 XX
 PD 17-NOV-1998.
 XX
 PR 07-JUN-1995; 95US-0487113.
 PR 07-JUN-1995; 95US-0487113.
 PR 07-JUN-1995; 95US-0487113.


```

PR 27-JAN-1992; 92US-0827689.
PR 26-MAY-1992; 92US-0889724.
PR 05-JUN-1992; 92US-0894061.
PR 22-JAN-1993; 93US-0009266.
PR 26-JAN-1993; 93US-0009266.
PR 05-AUG-1993; 93US-0102852.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX MPI: 1999-023535/02.
XX N-PSDB; AAV69185.
XX
XX Humanised antibodies specific for intercellular adhesion molecule
XX polypeptide - useful for therapeutic or diagnostic purposes
XX
XX Example 13; Columns 137-138; 116pp; English.
XX
XX The invention relates to humanised ICR-1.1 and ICR-8.1 antibodies
XX targeted to the human intercellular adhesion molecule polypeptide
XX (ICAM-R) polypeptide. Antibodies specific for ICAM-R's are potentially
XX useful as therapeutic compounds, for treating e.g. immune-mediated
XX inflammatory conditions (e.g. graft-versus-host disease) asthma,
XX tumours or viral infections. Monoclonal antibodies specific for ICAM-R,
XX or their conjugates formed with e.g. toxins or radionuclides are useful
XX for therapeutically targeting or detecting neovascularisation sites.
XX The present sequence represents the amino acid sequence of the Vh region
XX of the murine antibody ICR-1.1.
XX
XX Sequence 138 AA:
SQ
Query Match 78.8%; Score 78; DB 20; Length 138;
Best Local Similarity 76.5%; Pred. No. 0.00015;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 1 YINPRNDYTEYNQNFKD 17
Db 69 YINPNTDTEYNQRFOD 85
RESULT 12
AAB13047
ID AAB13047 standard; Protein; 138 AA.
XX
XX AAB13047;
AC
XX
XX 19-DEC-2000 (first entry)
DT
XX
XX Murine ICR-1.1 Vh protein sequence.
DE
XX
XX Anti-human immunodeficiency virus; HIV; cytostatic; ICAM-R; ARDS; stroke;
XX intercellular adhesion molecule; immunoglobulin heavy chain; septicemia;
XX inflammatory conditions; glomerulonephritis; arthritis; dermatosis;
XX haemodialysis; leukapheresis; ulcerative colitis; Crohn's disease;
XX necrotising enterocolitis; atherosclerosis; psoriasis; asthma;
XX transplant rejection; diabetes; tumour.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 2
FT Misc-difference /Label- Unknown
FT Misc-difference 6
FT Misc-difference /Label- Unknown
FT Misc-difference 7
FT Misc-difference /Label- Unknown
FT Misc-difference 8
FT Misc-difference /Label- Unknown
XX
XX US6100383-A.
XX
XX 08-AUG-2000.

```

```

XX
XX 07-JUN-1995; 95US-0475680.
XX
XX 05-AUG-1994; 94US-0286754.
XX 26-JAN-1993; 93US-0009266.
XX 27-JAN-1992; 92US-0827689.
XX 26-MAY-1992; 92US-0889724.
XX 05-JUN-1992; 92US-0894061.
XX 22-JAN-1993; 93US-0009266.
XX 05-AUG-1993; 93US-0102852.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX MPI: 2000-542449/49.
XX N-PSDB; AAV97155.
XX
XX Hybrid fusion proteins comprising intercellular adhesion molecule or
XX its variants useful, for treating inflammatory conditions, Crohn's
XX disease, atherosclerosis and diabetes
XX
XX Example 13; Column 127-128; 109pp; English.
XX
XX This invention relates to a hybrid fusion protein comprising an
XX intercellular adhesion molecule (ICAM-R) amino acid fragment at its
XX amino terminus and a constant domain of an immunoglobulin heavy chain at
XX its carboxy terminus. ICAM-R polypeptides are useful for treating and
XX monitoring inflammatory conditions such as adult respiratory distress
XX syndrome, multiple organ injury syndrome secondary to septicemia or
XX trauma, reperfusion injury of tissue, acute glomerulonephritis, reactive
XX arthritis, dermatosis, stroke, thermal injury, haemodialysis,
XX leukapheresis, ulcerative colitis, Crohn's disease, necrotising
XX enterocolitis, granulocyte transfusion associated syndrome,
XX atherosclerosis and cytokine induced toxicity. ICAM-R polypeptides are
XX also useful for treating conditions resulting from a response of the
XX specific immune system in a mammal e.g. psoriasis, organ/tissue
XX transplant rejection and autoimmune diseases including Raynaud's
XX syndrome, autoimmune thyroiditis, multiple sclerosis, rheumatoid
XX arthritis, diabetes and lupus erythematosus. ICAM-R products and ICAM-R
XX related products are also useful in monitoring and treating asthma,
XX tumour growth and/or metastasis, and viral infection (e.g. HIV
XX infection). Sequences AAB97090 and AAB13036 represent the human ICAM-R
XX DNA and protein sequences. Sequences AAB97091-A97112 represent ICAM-R
XX DNA fragments, PCR primers and probes, all used in the identification of
XX the ICAM-R DNA sequence. AAB97113-A97123 and AAB97129-A97152 represent
XX primers used in the production of humanised anti-ICAM-R antibody ICR-8.1,
XX and fragments of the humanised antibody. Sequences AAB97124-A97128,
XX AAB97132, AAB97144 represent ICR-8.1 sequences. Sequences AAB97153-A97176
XX excluding AAB97155-A97156 represent primers used in the production of
XX humanised anti-ICAM-R antibody ICR-1.1, and fragments of the humanised
XX antibody. Sequences AAB97155-A97156 and AAB13047-B13048 represent murine
XX ICR-1.1 sequences. DNA and peptide sequences used in the production of
XX the chimeric protein of the invention include AAB97177-A97188 and
XX AAB13050-B13051.
XX
XX Sequence 138 AA:
SQ
Query Match 78.8%; Score 78; DB 21; Length 138;
Best Local Similarity 76.5%; Pred. No. 0.00015;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 1 YINPRNDYTEYNQNFKD 17
Db 69 YINPNTDTEYNQRFOD 85
RESULT 13
AAV82446
ID AAV82446 standard; Protein; 138 AA.
XX
XX AAV82446;
AC
XX

```

DT 28-JUN-2000 (first entry)
XX Murine ICR-1.1 antibody VH protein sequence SEQ ID NO:78.
XX
XX Human; ICAM-R; chromosome 19; intracellular adhesion molecule receptor;
KM CAM; ICAM-1; ICAM-2; humanised; antibody; mutagenic; chimeric; vulnarary;
KM nephropathic; antirheumatic; cerebroprotective; antitumor; cytotoxic;
KM antileukemic; antitumor; immunosuppressive; antidiabetic; neuroprotective;
KM antihypertensive; dermatological; antitumor; antitumor; antitumor;
KM anti-HIV; vasotropic; antipsoriatic; immunomodulatory; antineoplastic;
KM cell adhesion mediator; inflammatory condition; immunisation;
KM immune response.
XX
XX Mus sp.
XX
XX US6040176-A.
XX
XX 21-MAR-2000.
XX
XX 12-SEP-1996; 96US-0714017.
XX
XX 05-AUG-1994; 94US-0286754.
XX 27-JAN-1992; 92US-0827689.
XX 26-MAY-1992; 92US-0889724.
XX 05-JUN-1992; 92US-0894061.
XX 22-JAN-1993; 93US-0009266.
XX 26-JAN-1993; 93US-0050787.
XX 05-AUG-1993; 93US-0102852.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX WPI; 2000-270138/23.
XX
XX Novel monoclonal antibody directed against ICAM-R proteins useful for
PT treating acute glomerulonephritis, ulcerative colitis, psoriasis
PT rheumatoid arthritis, diabetes, multiple sclerosis, asthma and viral
PT infection -
XX
XX Example 13; Column 139; 117pp; English.
XX
XX The present invention describes a monoclonal antibody (MAb) (I),
CC produced by the hybridoma cell line 81K2F (ATCC HB 11692). Also described
CC are: (1) a hybridoma cell line 81K2F; and (2) a MAb (II), that competes
CC with (I) for binding to ICAM-R (intracellular adhesion molecule
CC receptor) (III). (II) mimics the activity of natural binding proteins
CC through which intercellular and intracellular activities of (III) are
CC modulated. (II) is also used for modulating the immune responses. (I) is
CC used for immunisation as well as for purifying (III). They are also
CC useful in modulating the ligand/receptor binding biological activity
CC involving (III) especially those effector functions of (III) involved in
CC specific and non-specific immune system responses. Inflammatory
CC conditions which may be treated or monitored with related products of
CC (III) include conditions resulting from a response of the non-specific
CC immune system in a mammal e.g. adult respiratory distress syndrome,
CC multiple organ injury syndrome secondary to septicemia or trauma,
CC reperfusion injury of tissue, acute glomerulonephritis, reactive
CC arthritis, stroke, ulcerative colitis and atherosclerosis, and conditions
CC resulting from a response of the specific immune system in a mammal, e.g.
CC psoriasis, organ/tissue transplantation rejection, autoimmune diseases
CC such as autoimmune thyroiditis, multiple sclerosis, rheumatoid arthritis,
CC diabetes and lupus erythematosus. AAB0236 to AAB0834, and AAB2435 to
CC AAB8251 represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 138 AA:
XX
XX Query Match 78.88; Score 78; DB 21; Length 138;
XX Best Local Similarity 76.5%; Pred. No. 0.00015;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DB 69 YINPNTDYTYNORFOD 85
||||| ||||| ||:|
RESULT 14
AAV50754
ID AAV50754 standard; Protein: 138 AA.
XX
XX AAV50754;
AC
XX 16-FEB-2000 (first entry)
XX
XX Murine antibody ICR-1.1 VH region protein.
XX
XX
XX ICAM-R; murine; intercellular adhesion molecule; phosphorylation;
KM protein kinase C; modulator; ICR-1.1; antibody.
XX
XX Mus sp.
XX
XX US5989843-A.
XX
XX 23-NOV-1999.
XX
XX 27-SEP-1996; 96US-0720420.
XX
XX 27-JAN-1992; 92US-0827689.
XX 26-MAY-1992; 92US-0889724.
XX 05-JUN-1992; 92US-0894061.
XX 22-JAN-1993; 93US-0009266.
XX 26-JAN-1993; 93US-0050787.
XX 05-AUG-1993; 93US-0102852.
XX 07-JUN-1993; 93US-0487113.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Vazeux R;
XX
XX WPI; 2000-022778/02.
XX
XX N-PSDB; AAZ24327.
XX
XX Identifying modulators of protein kinase C phosphorylation of human
PT intercellular adhesion molecule polypeptide -
XX
XX Example 14; Column 143-144; 122pp; English.
XX
XX This invention describes a novel method for identifying a compound that
CC modulates phosphorylation of human intercellular adhesion molecule
CC polypeptide (ICAM-R) by protein kinase C isoform. The method comprises:
CC (a) exposing a purified peptide consisting of the cytoplasmic domain of
CC ICAM-R to protein kinase C isoform and labeled adenosine triphosphate in
CC the presence and absence of a test compound; (b) measuring labeled
CC phosphate transferred to the peptide; and (c) identifying a test compound
CC that affects transfer of the labeled phosphate as a modulator compound.
CC The method is useful for identifying compounds that modulate the
CC phosphorylation of human intercellular adhesion molecule polypeptide
CC which might form the basis for the development of therapeutic and
CC diagnostic agents. This sequence represents the murine ICR-1.1 antibody
CC VH region which is used in the method of the invention.
XX
XX Sequence 138 AA:
XX
XX Query Match 78.88; Score 78; DB 21; Length 138;
XX Best Local Similarity 76.5%; Pred. No. 0.00015;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 YINPNDYTYNORFOD 17
||||| ||||| ||:|
RESULT 15
AAU70939
ID AAU70939 standard; Protein: 138 AA.

XX AC AUA70939;
 XX DT 26-FEB-2002 (first entry)
 XX DE Murine monoclonal antibody ICR-1.1 VH region.
 XX
 KW Human; intercellular adhesion molecule; ICAM; antiinflammatory; stroke;
 KW antibacterial; vulnary; vasotropic; nephrotoxic; antiarthritic;
 KW cerebroprotective; dermatological; antiulcer; immunosuppressive; tumour;
 KW antiproliferative; antiarteriosclerotic; neuroprotective; antithyroid;
 KW vitruclide; antirheumatic; antidiabetic; antistatic; cytosolic; asthma;
 KW hybridoma cell line; ATCC HB 12190; inflammation; septicemia; trauma;
 KW adult respiratory distress syndrome; multiple organ injury syndrome;
 KW tissue reperfusion injury; acute glomerulonephritis; arthritis; vaccine;
 KW dermatosis; thermal injury; haemodialysis; leukoparesis; psoriasis;
 KW Crohn's disease; ulcerative colitis; multiple sclerosis; infection.
 XX
 OS Mus sp.
 XX
 PN US2001029293-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 03-JAN-2001; 2001US-0753436.
 XX
 PR 24-AUG-1999; 99US-0382289.
 PR 27-JAN-1992; 92US-0827689.
 PR 26-MAY-1992; 92US-0889724.
 PR 05-JUN-1992; 92US-0894061.
 PR 22-JAN-1993; 93US-0009266.
 PR 26-JAN-1993; 93WO-0500787.
 PR 05-AUG-1993; 93US-0102852.
 PR 07-JUN-1995; 95US-0487113.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 P1 Gallatin WM, Vazeux R;
 XX
 DR WPI; 2002-009992/01.
 DR P-PSDB; AUA70939.
 XX
 PT Novel hybridoma cell line useful for producing monoclonal antibody for
 PT treating inflammatory conditions, immune system disorders and
 PT infectious diseases, is deposited under specified ATCC accession number
 PT
 PS Page 78; Example 14; 126pp; English.
 XX
 CC The invention relates to a novel hybridoma cell line (I) ATCC HB 12190.
 CC (I) is useful for producing an intercellular adhesion molecule (ICAM)
 CC monoclonal antibody (II). (II) is useful for treating inflammatory
 CC conditions including adult respiratory distress syndrome, multiple organ
 CC injury syndrome secondary to septicemia or trauma, tissue reperfusion
 CC injury, acute glomerulonephritis, reactive arthritis, dermatosis with
 CC acute inflammatory components, stroke, thermal injury, haemodialysis,
 CC leukoparesis, ulcerative colitis, Crohn's disease, necrotising
 CC enterocolitis, granulocyte transfusion associated syndrome, diabetes,
 CC atherosclerosis, cytokine-induced toxicity, psoriasis, organ/tissue
 CC transplant rejection, autoimmune diseases including Raynaud's syndrome,
 CC autoimmune thyroiditis, multiple sclerosis, rheumatoid arthritis,
 CC lupus erythematosus, asthma, tumour growth and/or metastasis, viral
 CC infection, tissue transplant rejection, graft versus host disease and
 CC multiple sclerosis. (II) is also useful for immunisation, for purifying
 CC ICAM-R polypeptides and for identifying cells that display the
 CC polypeptides on their surfaces. AUA70928-AUA70946 represent ICAM
 CC amino acid sequences of the invention.
 CC
 SQ Sequence 138 AA;

OY 1 YINPRNDYEXNQNPKD 17
 |||| ||||| :|
 DB 69 YINPNTDYEXNQRQD 85

Search completed: October 7, 2003, 19:13:57
 Job time : 21.3966 secs

Query Match 78.8%; Score 78; DB 23; Length 138;
 Best Local Similarity 76.5%; Pred. No. 0.00015;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 ; Search time 8.7931 Seconds
(without alignments)
305.878 Million cell updates/sec

Title: US-09-988-013A-4_COPY_50_66

Perfect score: 99
Sequence: 1 YINPRNDYTEYNONFKD 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications, AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	116	10	US-09-741-843-4
2	99	100.0	116	10	US-09-741-843-8
3	99	100.0	116	10	US-09-741-843-9
4	99	100.0	116	11	US-09-894-839-4
5	99	100.0	116	11	US-09-894-839-8
6	99	100.0	116	11	US-09-894-839-21
7	99	100.0	116	11	US-09-888-013A-4
8	99	100.0	116	11	US-09-888-013A-8
9	99	100.0	116	11	US-09-888-013A-9
10	99	100.0	116	11	US-09-888-013A-9
11	99	100.0	116	12	US-10-127-890-167
12	99	100.0	116	12	US-10-127-890-167
13	99	100.0	116	12	US-10-127-890-169
14	99	100.0	116	12	US-10-244-821-91
15	99	100.0	116	12	US-10-244-821-91

16	70	70.7	256	15	US-10-247-488-2	Sequence 2, Appl1
17	70	70.7	258	15	US-10-247-488-4	Sequence 4, Appl1
18	69	69.7	119	11	US-09-795-515-11	Sequence 11, Appl1
19	69	69.7	119	11	US-09-795-515-30	Sequence 30, Appl1
20	69	69.7	119	15	US-10-267-286A-10	Sequence 10, Appl1
21	69	69.7	119	15	US-10-267-286A-12	Sequence 12, Appl1
22	69	69.7	119	15	US-10-267-286A-13	Sequence 13, Appl1
23	69	69.7	119	15	US-10-267-286A-14	Sequence 14, Appl1
24	69	69.7	468	11	US-09-795-515-7	Sequence 7, Appl1
25	69	69.7	468	11	US-09-795-515-7	Sequence 11, Appl1
26	68	68.7	116	12	US-10-168-809-22	Sequence 22, Appl1
27	65	65.7	116	14	US-10-127-890-168	Sequence 168, App
28	65	65.7	104	14	US-10-032-482-1	Sequence 9, Appl1
29	65	65.7	127	12	US-09-998-831-7	Sequence 1, Appl1
30	65	65.7	127	12	US-09-998-831-7	Sequence 7, Appl1
31	64	64.6	195	12	US-10-360-053-6	Sequence 6, Appl1
32	64	64.6	195	12	US-10-360-053-8	Sequence 8, Appl1
33	63	63.6	17	15	US-10-283-349-41	Sequence 41, Appl1
34	63	63.6	140	15	US-10-283-349-27	Sequence 27, Appl1
35	63	63.6	140	15	US-10-283-349-63	Sequence 63, Appl1
36	63	63.6	140	15	US-10-283-349-74	Sequence 74, Appl1
37	63	63.6	140	15	US-10-283-349-83	Sequence 83, Appl1
38	63	63.6	140	15	US-10-283-349-83	Sequence 83, Appl1
39	63	63.6	242	16	US-10-259-087A-20	Sequence 20, Appl1
40	62	62.6	119	11	US-09-795-515-12	Sequence 12, Appl1
41	62	62.6	119	11	US-09-795-515-13	Sequence 13, Appl1
42	62	62.6	119	11	US-09-795-515-14	Sequence 14, Appl1
43	62	62.6	119	11	US-09-795-515-15	Sequence 15, Appl1
44	62	62.6	119	11	US-09-795-515-16	Sequence 16, Appl1
45	62	62.6	119	11	US-09-795-515-17	Sequence 17, Appl1

ALIGNMENTS

RESULT 1
US-09-741-843-4
; Sequence 4, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hens
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-741-843-4
Query Match 100.0%; Score 99; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFKD 17
DB 50 YINPRNDYTEYNONFKD 66
RESULT 2
US-09-741-843-8
; Sequence 8, Application US/09741843

```
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYM
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-8
```

```
Query Match          100.0%; Score 99; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 9,8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 YINPRNDYTEYNONFKD 17
         |||||||
Db       50 YINPRNDYTEYNONFKD 66
```

```
RESULT 3
US-09-741-843-9
; Sequence 9, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYM
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-9
```

```
Query Match          100.0%; Score 99; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 9,8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 YINPRNDYTEYNONFKD 17
         |||||||
Db       50 YINPRNDYTEYNONFKD 66
```

```
RESULT 4
US-09-894-839-4
; Sequence 4, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
```

```
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-4
```

```
Query Match          100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9,8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 YINPRNDYTEYNONFKD 17
         |||||||
Db       50 YINPRNDYTEYNONFKD 66
```

```
RESULT 5
US-09-894-839-8
; Sequence 8, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-8
```

```
Query Match          100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9,8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 YINPRNDYTEYNONFKD 17
         |||||||
Db       50 YINPRNDYTEYNONFKD 66
```

```
RESULT 6
US-09-894-839-21
; Sequence 21, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
```

;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: US 60/013,709
;; PRIOR FILING DATE: 1996-03-20
;; NUMBER OF SEQ ID NOS: 47
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 21
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-894-839-21

Query Match 100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 7
US-09-988-013A-4
;; Sequence 4, Application US/09988013A
;; Publication No. US20030103979A1
;; GENERAL INFORMATION:
;; APPLICANT: LEUNG, Shui-on
;; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
;; FILE REFERENCE: 18733/1082
;; CURRENT APPLICATION NUMBER: US/09/988,013A
;; PRIOR FILING DATE: 2002-10-29
;; PRIOR APPLICATION NUMBER: US 09/741,843
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/127,902
;; PRIOR FILING DATE: 1998-08-03
;; PRIOR APPLICATION NUMBER: US 08/690,102
;; PRIOR FILING DATE: 1996-07-06
;; PRIOR APPLICATION NUMBER: US 08/289,576
;; PRIOR FILING DATE: 1994-08-12
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 4
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Murinae gen. sp.
US-09-988-013A-4

Query Match 100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 8
US-09-988-013A-8
;; Sequence 8, Application US/09988013A
;; Publication No. US20030103979A1
;; GENERAL INFORMATION:
;; APPLICANT: LEUNG, Shui-on
;; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
;; FILE REFERENCE: 18733/1082
;; CURRENT APPLICATION NUMBER: US/09/988,013A
;; PRIOR FILING DATE: 2002-10-29
;; PRIOR APPLICATION NUMBER: US 09/741,843
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/127,902
;; PRIOR FILING DATE: 1998-08-03

;; PRIOR APPLICATION NUMBER: US 08/690,102
;; PRIOR FILING DATE: 1996-07-06
;; PRIOR APPLICATION NUMBER: US 08/289,576
;; PRIOR FILING DATE: 1994-08-12
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-988-013A-8

Query Match 100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 9
US-09-988-013A-9
;; Sequence 9, Application US/09988013A
;; Publication No. US20030103979A1
;; GENERAL INFORMATION:
;; APPLICANT: LEUNG, Shui-on
;; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
;; FILE REFERENCE: 18733/1082
;; CURRENT APPLICATION NUMBER: US/09/988,013A
;; PRIOR FILING DATE: 2002-10-29
;; PRIOR APPLICATION NUMBER: US 09/741,843
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/127,902
;; PRIOR FILING DATE: 1998-08-03
;; PRIOR APPLICATION NUMBER: US 08/690,102
;; PRIOR FILING DATE: 1996-07-06
;; PRIOR APPLICATION NUMBER: US 08/289,576
;; PRIOR FILING DATE: 1994-08-12
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 9
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-988-013A-9

Query Match 100.0%; Score 99; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17
DB 50 YINPRNDYTEYNQNFKD 66

RESULT 10
US-09-753-436-78
;; Sequence 78, Application US/09753436
;; Patent No. US20010029293A1
;; GENERAL INFORMATION:
;; APPLICANT: Gallatin, W. Michael
;; TITLE OF INVENTION: ICAM-Related Materials and Methods
;; NUMBER OF SEQUENCES: 120
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: United States of America

ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/753,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/382,289
FILING DATE:
APPLICATION NUMBER: US 08/487,113
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,754
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Joseph A., Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 33282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6500
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-753-436-78

Query Match 78.88; Score 78; DB 9; Length 138;
Best Local Similarity 76.58; Pred. No. 0.00018;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 YINPRNDYTEYNOKFD 17
||||| ||||| |||
Db 69 YINPRNDYTEYNOKFD 85

RESULT 11
US-10-127-890-167
Sequence 167, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago

STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 167:
US-10-127-890-167

Query Match 74.78; Score 74; DB 12; Length 116;
Best Local Similarity 76.58; Pred. No. 0.00059;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 YINPRNDYTEYNOKFD 17
||||| ||||| |||
Db 50 YINPRNDYTEYNOKFD 66

RESULT 12
US-10-127-890-169
Sequence 169, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible


```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-10-127-890-169

Query Match      74.7%; Score 74; DB 12; Length 116;
Best Local Similarity 76.5%; Pred. No. 0.00059;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy      1 YINPRNDYTEYNQKFD 17
      ||| ||||| |||
Db      50 YINPSTGYTEYNQKFD 66

RESULT 13
US-10-244-821-91
; Sequence 91, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-244-821-91

Query Match      74.7%; Score 74; DB 12; Length 135;
Best Local Similarity 76.5%; Pred. No. 0.0007;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy      1 YINPRNDYTEYNQKFD 17
      ||| ||||| |||
Db      69 YINPSTGYTEYNQKFD 85

RESULT 14
US-10-244-821-88
; Sequence 88, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-244-821-88

Query Match      74.7%; Score 74; DB 12; Length 438;
Best Local Similarity 76.5%; Pred. No. 0.0025;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy      1 YINPRNDYTEYNQKFD 17
      ||| ||||| |||
Db      72 YINPSTGYTEYNQKFD 88

RESULT 15
US-09-753-436-86
; Sequence 86, Application US/09753436
; Patent No. US20010029293A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,436
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/382,289
; FILING DATE:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Joseph A., Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 33282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6500
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-753-436-86

Query Match 72.78; Score 72; DB 9; Length 123;
Best Local Similarity 75.08; Pred. No. 0.0013;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNQNFK 16
||| ||||| |
Db 54 YINPNTDYTEYNQRFQ 69

Search completed: October 7, 2003, 19:24:27
Job time: 8.7931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 6.33103 Seconds

(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013a-4_COPY_50_66

Perfect score: 99

Sequence: 1 YINPRNDYTEYNQNFKD 17

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	81.8	139	2 PS0024	Ig heavy chain pre
2	80	80.8	96	2 SI7615	Ig heavy chain v r
3	75	75.8	102	2 S26471	Ig heavy chain v r
4	75	75.8	106	2 PH1005	Ig heavy chain v r
5	75	75.8	139	2 A27609	Ig heavy chain pre
6	74	74.7	96	2 SI7612	Ig heavy chain v r
7	74	74.7	96	2 SI7621	Ig heavy chain v r
8	74	74.7	96	2 SI7606	Ig heavy chain v r
9	74	74.7	96	2 SI7616	Ig heavy chain v r
10	74	74.7	96	2 SI7602	Ig heavy chain v r
11	74	74.7	96	2 SI7611	Ig heavy chain v r
12	74	74.7	96	2 SI7619	Ig heavy chain v r
13	74	74.7	96	2 SI7607	Ig heavy chain v r
14	74	74.7	96	2 SI7613	Ig heavy chain v r
15	74	74.7	96	2 SI7617	Ig heavy chain v r
16	74	74.7	84	2 PH1487	Ig heavy chain v r
17	70	70.7	96	2 SI7230	Ig heavy chain v r
18	70	70.7	96	2 SI7618	Ig heavy chain v r
19	70	70.7	96	2 SI7620	Ig heavy chain v r
20	70	70.7	96	2 SI7614	Ig heavy chain v r
21	70	70.7	96	2 SI7610	Ig heavy chain v r
22	70	70.7	121	2 A21854	Ig heavy chain v r
23	69	69.7	98	2 B24754	Ig heavy chain v r
24	67	67.7	114	2 PL0247	Ig heavy chain v r
25	67	67.7	117	2 JC2269	PL7-6 antibody hea
26	67	67.7	123	2 G48677	Ig heavy chain v-D
27	66	66.7	98	2 A28572	Ig heavy chain v r
28	66	66.7	114	2 PH1522	Ig heavy chain v r
29	66	66.7	114	2 PH1523	Ig heavy chain v r

30	66	66.7	118	2 S38565	Ig heavy chain v r
31	66	66.7	119	2 PH1517	Ig heavy chain v r
32	66	66.7	119	2 PH1521	Ig heavy chain v r
33	66	66.7	119	2 PH1504	Ig heavy chain v r
34	66	66.7	119	2 PH1500	Ig heavy chain v r
35	66	66.7	119	2 PH1502	Ig heavy chain v r
36	66	66.7	119	2 PH1505	Ig heavy chain v r
37	66	66.7	119	2 PH1512	Ig heavy chain v r
38	66	66.7	121	2 A26405	Ig heavy chain v r
39	66	66.7	140	2 PH1482	Ig heavy chain v r
40	66	66.7	140	2 PH1484	Ig heavy chain v r
41	66	66.7	140	2 S26320	Ig heavy chain v r
42	64	64.6	107	2 PH0996	Ig heavy chain v r
43	63	63.6	94	2 SI7603	Ig heavy chain v r
44	63	63.6	97	2 SI7603	Ig heavy chain v r
45	63	63.6	98	2 S26313	Ig heavy chain v r

ALIGNMENTS

RESULT 1

PS0024

Ig heavy chain precursor V region (6A4) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 22-Nov-1996

C:Accession: PS0024

R:Margel, M.; Eckhardt, A.; Ehret, W.; von Specht, B.U.; Duchene, M.; Domdey, H.

Gene 74, 335-345, 1988

A:Title: Cloning and characterization of cDNAs coding for the heavy and light chains

A:Reference number: PS0023; MUID:85232725; PMID:314944

A:Accession: PS0024

A:Molecule type: mRNA

A:Residues: 1-139 <MAR>

A:Experimental source: strain BALB/c

C:Comment: This chain is obtained from an IgG2a monoclonal antibody against Pseudomon

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin; pyroglyutamic acid

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-139/Domain: Ig heavy chain V region #status predicted <IGV>

F:34-117/Domain: immunoglobulin homology <IMH>

F:20/Modified site: pyroglutamic acid (Gln) (in mature form) #status predic

Query Match

Best Local Similarity 81.8%; Score 81; DB 2; Length 139;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFKD 17

DB 69 YINPRNDYTEYNQNFKD 85

RESULT 2

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: SI7615

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: SI7230; MUID:91326098; PMID:1907718

A:Accession: SI7615

A:Molecule type: nucleic acid

A:Status: preliminary

A:Residues: 1-96 <CIA>

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMH>

Query Match

Best Local Similarity 80.8%; Score 80; DB 2; Length 96;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNONFKD 17
||||| ||||| |||
Db 43 YINPRTGYTEYNOKFKD 59

RESULT 3

S26471
Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_rev1sion 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26471
R:Kavaler, J.
Submitted to the EMBL Data Library, April 1991
A:Reference number: S26459
A:Accession: S26471
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-102 <RAV>
A:Cross-references: EMBL:X59110; NID:g51951; PIDN:CAA41836.1; PID:g51952
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 75.8%; Score 75; DB 2; Length 102;
Best Local Similarity 76.5%; Pred. No. 0.00015;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFKD 17
||||| ||||| |||
Db 50 YINPSSGYTEYNOKFKD 66

RESULT 4

PH1005
Ig heavy chain V region (clone 202.54) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_rev1sion 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PH1005
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B cells
A:Reference number: PH0971; MUID:92381444; PMID:1512540
A:Accession: PH1005
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-106 <RIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 75.8%; Score 75; DB 2; Length 106;
Best Local Similarity 76.5%; Pred. No. 0.00015;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFKD 17
||||| ||||| |||
Db 50 YINPSSGYTEYNOKFKD 66

RESULT 5

A27609
Ig heavy chain precursor V region (129) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 15-Dec-1988 #sequence_rev1sion 30-Jun-1991 #text_change 23-Jul-1999
C:Accession: A27609
R:Klein, D.; Nieuwkoop, J.; Sirin, S.; Stavnezer, J.
J. Immunol. 140, 1676-1684, 1988
A:Title: 1.29 lymphoma cells express a nonmutated V-H gene before and after H chain switch
A:Reference number: A27609; MUID:88154467; PMID:3126234
A:Accession: A27609
A:Molecule type: DNA

A:Residues: 1-139 <KLE>
A:Cross-references: EMBL:M19401; NID:g195441; PIDN:AAA38303.1; PID:g553992
C:Genetics: 16/1
A:Insertions: 16/1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-139/Product: Ig heavy chain V region 129 #status predicted <VAR>
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 75.8%; Score 75; DB 2; Length 139;
Best Local Similarity 81.2%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFK 16
||||| ||||| |||
Db 69 YINPRNDYTEYNOKFK 84

RESULT 6

S17612
Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Nov-1993 #sequence_rev1sion 10-Nov-1995 #text_change 06-Jun-1997
C:Accession: S17612
R:Jackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.
Nature 352, 624-628, 1991
A:Title: Making antibody fragments using phage display libraries.
A:Reference number: S17230; MUID:91326098; PMID:1907718
A:Accession: S17612
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-96 <CLA>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin
F:8-91/Domain: Immunoglobulin homology <IMM>

Query Match 74.7%; Score 74; DB 2; Length 96;
Best Local Similarity 76.5%; Pred. No. 0.0002;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFKD 17
||||| ||||| |||
Db 43 YINPRTGYTEYNOKFKD 59

RESULT 7

S17621
Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Nov-1993 #sequence_rev1sion 10-Nov-1995 #text_change 06-Jun-1997
C:Accession: S17621
R:Jackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.
Nature 352, 624-628, 1991
A:Title: Making antibody fragments using phage display libraries.
A:Reference number: S17230; MUID:91326098; PMID:1907718
A:Accession: S17621
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-96 <CLA>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; Immunoglobulin
F:8-91/Domain: Immunoglobulin homology <IMM>

Query Match 74.7%; Score 74; DB 2; Length 96;
Best Local Similarity 76.5%; Pred. No. 0.0002;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 YINPRNDYTEYNONFKD 17
||||| ||||| |||
Db 43 YINPRTGYTEYNOKFKD 59

RESULT 8

S17606

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17606

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17606

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <C1A>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.7%; Score 74; DB 2; Length 96;

Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFKD 17

DB 43 YINPSTGYTEYNQFKD 59

RESULT 9

S17616

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17616

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17616

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <C1A>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.7%; Score 74; DB 2; Length 96;

Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFKD 17

DB 43 YINPSTGYTEYNQFKD 59

RESULT 10

S17602

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17602

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17602

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <C1A>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.7%; Score 74; DB 2; Length 96;

Best Local Similarity 76.5%; Pred. No. 0.0002; Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFKD 17

DB 43 YINPSTGYTEYNQFKD 59

RESULT 11

S17611

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17611

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17611

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <C1A>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.7%; Score 74; DB 2; Length 96;

Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFKD 17

DB 43 YINPSTGYTEYNQFKD 59

RESULT 12

S17619

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17619

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17619

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <C1A>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:8-91/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.7%; Score 74; DB 2; Length 96;

Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFKD 17

DB 43 YINPSTGYTEYNQFKD 59

RESULT 13

S17607

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S17607

R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.

Nature 352, 624-628, 1991

A:Title: Making antibody fragments using phage display libraries.

A:Reference number: S17230; MUID:91326098; PMID:1907718

A:Accession: S17607

A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-96 <CIA>
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:8-91/Domain: Immunoglobulin homology <IMM>

Query Match 74.7%; Score 74; DB 2; Length 96;
 Best Local Similarity 76.5%; Pred. No. 0.0002;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 YINPRNDYTEYNQNFKD 17
 ||||| ||||| |||||
 DB 43 YINPSTGYTEYNQNFKD 59

RESULT 14

S17613
 Ig heavy chain V region - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997
 C:Accession: S17613
 R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.
 Nature 352, 624-628, 1991
 A:Title: Making antibody fragments using phage display libraries.
 A:Reference number: S17230; MOID:91326098; PMID:1907718
 A:Accession: S17613
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-96 <CIA>
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:8-91/Domain: Immunoglobulin homology <IMM>

Query Match 74.7%; Score 74; DB 2; Length 96;
 Best Local Similarity 76.5%; Pred. No. 0.0002;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 YINPRNDYTEYNQNFKD 17
 ||||| ||||| |||||
 DB 43 YINPSTGYTEYNQNFKD 59

RESULT 15

S17617
 Ig heavy chain V region - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997
 C:Accession: S17617
 R:Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.
 Nature 352, 624-628, 1991
 A:Title: Making antibody fragments using phage display libraries.
 A:Reference number: S17230; MOID:91326098; PMID:1907718
 A:Accession: S17617
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-96 <CIA>
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:8-91/Domain: Immunoglobulin homology <IMM>

Query Match 74.7%; Score 74; DB 2; Length 96;
 Best Local Similarity 76.5%; Pred. No. 0.0002;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 YINPRNDYTEYNQNFKD 17
 ||||| ||||| |||||
 DB 43 YINPSTGYTEYNQNFKD 59

Search completed: October 7, 2003, 19:21:50
 Job time : 6.33103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 3.45862 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-4_COPY_50_66

Perfect score: 99

Sequence: 1 YINPRNDYTEYNONFKD 17

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.7	120	1	HY03_MOUSE	P01747 mus musculu
2	59.6	140	1	HY02_MOUSE	P01746 mus musculu
3	53.5	117	1	HY12_MOUSE	P01756 mus musculu
4	53.5	117	1	HY13_MOUSE	P01757 mus musculu
5	53.5	118	1	HY51_MOUSE	P06330 mus musculu
6	49.5	120	1	HY50_MOUSE	P06329 mus musculu
7	48.5	117	1	HY04_MOUSE	P01748 mus musculu
8	48.5	117	1	HY14_MOUSE	P01758 mus musculu
9	48.5	982	1	MSHM_SARGL	O63852 sarcophyton
10	47.5	48.0	284	HPRT_SCHMA	P03383 schistosoma
11	47.5	117	1	HY05_MOUSE	P01749 mus musculu
12	44.4	117	1	VNO1_VACCV	P17361 vaccinia vi
13	44.4	532	1	YHB7_YEAST	P28745 saccharomyc
14	44.4	1783	1	V468_MYCCE	O49460 mycoplasma
15	43.9	821	1	VOIG_ECOLI	P76655 escherichia
16	43.4	117	1	HY06_MOUSE	P01750 mus musculu
17	43.4	138	1	HY48_MOUSE	P03980 mus musculu
18	43.4	512	1	YE94_SCHPO	O13766 schizosacch
19	43.4	623	1	VGLE_VZVD	P09259 varicella-z
20	43.4	640	1	EX5A_HAETN	P45158 haemophilus
21	43.4	1168	1	MYSC_ACACA	P10569 acanthamoeb
22	42.5	398	1	YFCU_ECOLI	O43435 homo sapien
23	42.5	881	1	YFCU_ECOLI	P77196 escherichia
24	42.4	193	1	SODF_COXBU	P19685 coxiella bu
25	42.4	393	1	ILIS_CERAE	O28612 cercopithec
26	42.4	398	1	ILIS_HUMAN	P27930 homo sapien
27	42.4	430	1	PURA_LISTIN	O94795 listeria in
28	42.4	2515	1	TUD_DROME	P25823 drosophila
29	41.5	174	1	Y281_PYRHO	O58019 pyrococcus
30	41.5	41.9	704	TX37_CAEBL	O942c9 caenorhabdi
31	41.5	41.9	704	MSG2_YEAST	P13748 saccharomyc
32	41.4	117	1	HY09_MOUSE	P01753 mus musculu
33	41.4	117	1	HY10_MOUSE	P01754 mus musculu

34	41	41.4	117	1	HY49_MOUSE	P06328 mus musculu
35	41	41.4	139	1	HY07_MOUSE	P01751 mus musculu
36	41	41.4	232	1	RFAY_ECOLI	P27240 escherichia
37	41	41.4	271	1	PYR6_FREDI	P14880 fremyella d
38	41	41.4	325	1	HRCA_STAM	P45556 straphylococ
39	41	41.4	325	1	HRCA_STAM	O8wa8 staphylococ
40	41	41.4	331	1	SR49_CAEBL	O09212 caenorhabdi
41	41	41.4	650	1	PEX8_PICAN	O00925 pichia angu
42	41	41.4	704	1	Y590_METVA	O58010 methanococc
43	41	41.4	804	1	FPS_DROME	P18106 drosophilla
44	41	41.4	826	1	CRAA_BACUH	O94597 bacillus th
45	41	41.4	2547	1	PAFX_HUMAN	O93008 h probable

ALIGNMENTS

```
RESULT 1
HY03_MOUSE STANDARD; PRT; 120 AA.
ID HY03_MOUSE
AC P01747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 36-65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83131846; PubMed=6186498;
RA Siekevitz M., Gelfer M.L., Brodeur P., Riblet R.,
RA Marshak-Rothstein A.;
RT *The genetic basis of antibody production: the dominant anti-arsenate
RT idotype response of the strain A mouse.*;
RL Eur. J. Immunol. 12:1023-1032(1982).
CC -1- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER
CC DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE, THE AUTHORS
CC CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J
CC SEGMENT, JH2.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC HSSP; P01789; IMCP.
DR InterPro: IPR007110; Ig-1-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig-V.
DR Pfam: PF00047; Ig: 1.
DR SMART; SM00406; IgV: 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 111 IG-LIKE.
FT NON_TER 120 120
SQ SEQUENCE 120 AA: 13307 MW: FF04E4A167B654AF CRC64:

Query Match 66.7%; Score 66; DB 1; Length 120;
Best Local Similarity 68.8%; Pred. No. 0.0035;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 2
HY02_MOUSE STANDARD; PRT; 140 AA.
ID HY02_MOUSE
AC P01746;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 9367 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A/J;
RX MEDLINE=82152818; PubMed=6801765;
RA Sims J., Rabbits T.H., Estess P., Slaughter C., Tucker P.W.,
RA Capra J.D.;
RT "Somatic mutation in genes for the variable portion of the
RT Immunoglobulin heavy chain.";
RL Science 216:309-311(1982).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00493; AAA38128.1; -.
DR PIR: A94264; HVMSC7.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; HydrIdoma; Signal.
KM CHAIN 1 19
FT CHAIN 20 140 IG HEAVY CHAIN V REGION 9367.
FT DOMAIN 20 139 IG-LIKE.
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15514 MW; 25A4CB8E31DA5CE8 CRC64;

Query Match 59.6%; Score 59; DB 1; Length 140;
Best Local Similarity 62.5%; Pred. No. 0.045;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQNFK 16
   ||||| | | | | |
DB 69 YINPGNGTSTYNQFK 84

RESULT 3
HV12_MOUSE STANDARD; PRT; 117 AA.
ID HV12_MOUSE
AC P01756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region MOPC 104E.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE ASN-55.
RX MEDLINE=83075344; PubMed=6816276;
RA Khiry M.R., Fuhman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RA Hood L.E.;
RT "Complete amino acid sequence of a mouse mu chain: homology among
RT heavy chain constant region domains.";
RL Biochemistry 21:5415-5424(1982).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA
CC PROTEIN HAS ALSO BEEN DETERMINED.
CC -1- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
CC PIR: A02039; MHMS4E.
DR HSSP: P01789; IMCP.
DR InterPro: IPR007110; Ig_1like.

```

```

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KM Immunoglobulin V region; Glycoprotein.
FT DOMAIN 1 16
FT DISULFID 22 96 IG-LIKE.
FT CARBOHYD 55 55 BY SIMILARITY.
FT NON_TER 117 117 N-LINKED (GLCNAC. . .) (COMPLEX).
SQ SEQUENCE 117 AA; 12983 MW; 3CF8ACEBFE447E41 CRC64;

Query Match 53.5%; Score 53; DB 1; Length 117;
Best Local Similarity 66.7%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 INPRNDYTEYNQNFK 16
   ||| | | | | | |
DB 51 INPNGGTSTYNQFK 65

RESULT 4
HV13_MOUSE STANDARD; PRT; 117 AA.
ID HV13_MOUSE
AC P01757;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region J558.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=80078170; PubMed=6765983;
RA Schilling J., Clevinger B., Davie J.M., Hood L.;
RT "Amino acid sequence of homogeneous antibodies to dextran and DNA
RT rearrangements in heavy chain V-region gene segments.";
RL Nature 283:35-40(1980).
CC -1- MISCELLANEOUS: THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO
CC BIND DEXTRAN DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF
CC WHICH OCCUR IN THE D AND J SEGMENTS.
CC -1- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
CC PIR: A26242; MHMSJ5.
DR HSSP: P01789; IMCP.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KM Immunoglobulin V region.
FT DOMAIN 1 16 IG-LIKE.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13024 MW; 292E2AF4BE447E41 CRC64;

Query Match 53.5%; Score 53; DB 1; Length 117;
Best Local Similarity 66.7%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 INPRNDYTEYNQNFK 16
   ||| | | | | | |
DB 51 INPNGGTSTYNQFK 65

RESULT 5
HV51_MOUSE STANDARD; PRT; 118 AA.
ID HV51_MOUSE
AC P06350;
DT 01-JAN-1988 (Rel. 06, Created)

```



```

DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Ig heavy chain V region AC38 205.12.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=84182519; PubMed=6201362;
RA Dildrop R., Boyens J., Slekevitiz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes.";
RL EMO J. 3:517-523(1984).
DR PIR: A02040; MHMS38.
DR HSSP: P01789; IMCP.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 98 V SEGMENT.
FT DOMAIN 99 104 D SEGMENT.
FT DOMAIN 105 118 J SEGMENT.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 12934 MW; 94F7BEAC762A018 CRC64;

Query Match 53.5%; Score 53; DB 1; Length 118;
Best Local Similarity 66.7%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPRNDYTEYNONFK 16
DB 51 INPNGGTNYNEKFK 65

RESULT 6
HV50_MOUSE STANDARD; PRT; 120 AA.
ID HV50_MOUSE
AC P06329;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region AC38 15.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=84182519; PubMed=6201362;
RA Dildrop R., Boyens J., Slekevitiz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes.";
RL EMO J. 3:517-523(1984).
DR PIR: A02037; MHMS15.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 98 V SEGMENT.
FT DOMAIN 99 105 D SEGMENT.
FT DOMAIN 106 120 J SEGMENT.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13311 MW; 914453F426F09834 CRC64;

```

```

Query Match 49.5%; Score 49; DB 1; Length 120;
Best Local Similarity 60.0%; Pred. No. 1.2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPRNDYTEYNONFK 16
DB 51 INPNGGTNYNEKFK 65

RESULT 7
HV04_MOUSE STANDARD; PRT; 117 AA.
ID HV04_MOUSE
AC P01748;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 23 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Botchwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NpB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1 MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
DR PIR: A02030; HVMS23.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 23.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12772 MW; C530F829C906F69B CRC64;

Query Match 48.5%; Score 48; DB 1; Length 117;
Best Local Similarity 60.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPRNDYTEYNONFK 16
DB 70 INPNGGTNYNEKFK 84

RESULT 8
HV14_MOUSE STANDARD; PRT; 117 AA.
ID HV14_MOUSE
AC P01758;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 108A precursor.
GN IGH-VJ558.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81245215; PubMed=6789211;
RA Gliori D., Zakut R., Effron K., Rechavi G., Ram D., Cohen J.B.;
RT "Diversity of germ-line immunoglobulin VH genes."
RL Nature 292:426-430(1981).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00488; AAA38519.1; -
DR PIR: A02041; HVMS8A.
DR HSSP: P01810; 2FBJ.
DR MGD: MG1:96486; Igh-VJ558.
DR InterPro: IPR007110; Iq_1Ike.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003596; Iq_V.
DR Pfam: PF00047; Iq; 1.
DR SMART: SM00406; IqV; 1.
DR PROSITE: PS00835; Iq_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 108A.
FT DOMAIN 20 >117 IG-LIKE.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12972 MW; 428CB44DF25D1BC2 CRC64;

Query Match 48.5%; Score 48; DB 1; Length 117;
Best Local Similarity 62.5%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNQFK 16
   ||| | | | | | | |
DB 69 YIYPYNGTGYNQFK 84

RESULT 9
MSHM_SARGL
ID MSHM_SARGL STANDARD; PRT; 982 AA.
AC 063852;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Mitochondrial Muts protein homolog.
OS Sarcophyton glaucum.
OC Mitochondrion.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Alcyonacea;
OC Alcyonidae; Sarcophyton.
OX NCBI_TaxID=70919;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98210232; PubMed=9541536;
RA Pont-Kingdon G., Okada N.A., Macfarlane J.L., Beagley C.T.,
RA Watkins-Sims C.D., Cavalier-Smith T., Clark-Walker G.D.,
RA Wolstenholme D.R.;
RT "Mitochondrial DNA of the coral Sarcophyton glaucum contains a gene
RT for a homologue of bacterial Muts: a possible case of gene transfer
RT from the nucleus to the mitochondrion."
RL J. Mol. Evol. 46:419-431(1998).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF063191; AAC16386.1; -
DR InterPro: IPR002711; HNH.
DR InterPro: IPR003615; HNH_nuc.
DR InterPro: IPR000432; Muts_C.
DR InterPro: IPR002863; Muts_N.
DR Pfam: PF01844; HNH; 1.
DR Pfam: PF01624; Muts_I; 1.
DR Pfam: PF05192; Muts_III; 1.
DR Pfam: PF00488; Muts_V; 1.
DR ProDom: PD001263; Muts_C; 1.
DR SMART: SM00507; HNH; 1.
DR SMART: SM00534; Mutsac; 1.
DR SMART: SM00533; Mutsd; 1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR ATP-Binding; DNA-binding; Mitochondrion.
FT NP_BIND 698 705 ATP (POTENTIAL).
SQ SEQUENCE 982 AA; 111912 MW; A3F107319BC0E023 CRC64;

Query Match 48.5%; Score 48; DB 1; Length 982;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 INPRNDYTEYNQFK 17
   || | | | | | | | |
DB 377 LNPIIDYSELNRYKD 392

RESULT 10
HPRT_SCHMA
ID HPRT_SCHMA STANDARD; PRT; 284 AA.
AC P09383;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPR)
DE (HGPRase).
GN HGPR.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Puerto Rican;
RX MEDLINE=8830331; PubMed=3136439;
RA Craig S.P., McKerrow J.H., Newport G.N., Wang C.C.;
RT "Analysis of cDNA encoding the hypoxanthine-guanine
RT phosphoribosyltransferase (HGPRase) of Schistosoma mansoni; a
RT putative target for chemotherapy."
RL Nucleic Acids Res. 16:7087-7101(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Puerto Rican;
RX MEDLINE=89160320; PubMed=2701934;
RA Craig S.P., Muralidhar M.G., McKerrow J.H., Wang C.C.;
RT "Evidence for a class of very small introns in the gene for
RT hypoxanthine-guanine phosphoribosyltransferase in Schistosoma
RT mansoni."
RL Nucleic Acids Res. 17:1635-1647(1989).
CC -1- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: Purine salvage.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.

```


AC p38745;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 61.2 kDa protein in APW2-DUR3 Intergenic region
 DE precursor.
 GN YHL017W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Faveille A., Fulton L., Gattung S., Gelsel C., Kirsten J.,
 RA Kuchalla T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Lacroix P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaekis E., Vaughan K.,
 RA Vignelli D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII.";
 RL Science 265:2077-2082(1994).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Strong, to yeast PTM1 and some, to S.pombe
 CC SPAC6B5.07c.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U11583; AAB65070.1; -.
 DR PIR: S46831; S46831.
 DR SGD: S0001009; YHL017W.
 KW Hypothetical protein; Transmembrane; Signal.
 FT FT STGMVL 1 20
 FT CHAIN 21 532
 FT FT 194 214
 FT TRANSMEM 226 247
 FT TRANSMEM 263 283
 FT TRANSMEM 302 322
 FT TRANSMEM 331 351
 FT TRANSMEM 376 396
 FT TRANSMEM 416 436
 FT SEQUENCE 532 AA; 61205 MW; EF2CB7CB74006503 CRC64;
 QY Query Match 44.4%; Score 44; DB 1; Length 532;
 Best Local Similarity 63.6%; Pred. No. 30;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Db 4 PRNDYTEYKON 14
 480 PRNDYQEHNN 490
 RESULT 14
 Y468_MYCGE STANDARD; PRT; 1783 AA.
 AC Q49460;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG468.
 GN MG468.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 RN [2]
 RP REVISIONS.
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 879-985 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: SOME, TO MG064.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39728; AAC72488.1; -.
 DR EMBL: U01808; AAD12339.1; -.
 DR TIGR: MG468; -.
 DR InterPro: IPR003838; DUF214.
 DR Pfam: PF02687; FtsX; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT FT STGMVL 16 36
 FT TRANSMEM 917 937
 FT TRANSMEM 967 987
 FT TRANSMEM 1010 1030
 FT TRANSMEM 1084 1104
 FT TRANSMEM 1660 1680
 FT TRANSMEM 1709 1729
 FT TRANSMEM 1730 1750
 FT TRANSMEM 1752 1772
 FT SEQUENCE 1783 AA; 200168 MW; 87BD575AEC2E374B CRC64;
 QY Query Match 44.4%; Score 44; DB 1; Length 1783;
 Best Local Similarity 46.7%; Pred. No. 1e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Db 1 YINPRNDYTEYKONF 15
 1425 YIKPKWATRYSEKF 1439
 RESULT 15
 YQIG_ECOLI STANDARD; PRT; 821 AA.
 AC P76655; P77034; P77035;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

RESULT 2

09JUL75 PRELIMINARY; PRT; 109 AA.
ID 09JUL75
AC 09JUL75;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Anti-myosin immunoglobulin heavy chain variable region
DE (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=20448942; PubMed=10992488;
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.
RT "T-cell-dependent antibody response to the dominant epitope of
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
RT with cardiac myosin."
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL: AF206031; AAF69329.1; -.
DR HSSP: P01810; 2FBJ
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV_1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12118 MW; FF65E441BFF936A6 CRC64;

Query Match 63.6%; Score 63; DB 11; Length 109;
Best Local Similarity 68.8%; Pred. NO. 0.035;

Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YINPRNDYTEYNQNFK 16
Db 41 YINPYNDGTFYKNEFK 56

RESULT 3

09QXEX PRELIMINARY; PRT; 117 AA.
ID 09QXEX
AC 09QXEX;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Immunoglobulin heavy chain V-D-J region (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Clemens A., Rademakers A., Specht C., Koelsch E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ225174; CAB65237.1; -.
DR HSSP: P01810; 2FBJ
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV_1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13000 MW; CDD82AF84D49734 CRC64;

Query Match 53.5%; Score 53; DB 11; Length 117;
Best Local Similarity 66.7%; Pred. NO. 1.3;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPRNDYTEYNQNFK 16
Db 51 INPNGSTYQNFK 65

RESULT 4

09QXFO PRELIMINARY; PRT; 117 AA.
ID 09QXFO
AC 09QXFO;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Immunoglobulin heavy chain V-D-J region (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Clemens A., Rademakers A., Specht C., Koelsch E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ225171; CAB65236.1; -.
DR HSSP: P01789; IMCP.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV_1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13060 MW; D816AD0858A47EAC CRC64;

Query Match 53.5%; Score 53; DB 11; Length 117;
Best Local Similarity 66.7%; Pred. NO. 1.3;

Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPRNDYTEYNQNFK 16
Db 51 INPNGSTYQNFK 65

RESULT 5

091V67 PRELIMINARY; PRT; 143 AA.
ID 091V67
AC 091V67;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE V304-D-J-C mu protein (V304-D-J-C mu protein) (fragment).
GN V304-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6;
RC Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)acetyl (NP)."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB069912; BAB63928.1; -.
DR EMBL: AB069914; BAB63930.1; -.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.

DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 143
SQ SEQUENCE 143 AA; 15775 MW; 91BC6012B44EEBDF CRC64;
Query Match
Best Local Similarity 53.5%; Score 53; DB 11; Length 143;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 INPRNDYTEXNONFK 16
DB 51 IDPSDYTYNOKFK 65
RESULT 6
Q924P9 PRELIMINARY; PRT; 143 AA.
AC 0924P9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE V303-D-J-C mu protein (Fragment).
GN V303-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP)." ;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069916; BAB63932.1; -.
DR InterPro; IPR007110; I9-1like.
DR InterPro; IPR003006; I9_MHC.
DR Pfam; PF00047; I9; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 143
SQ SEQUENCE 143 AA; 15704 MW; 43CD8C72D52134F6 CRC64;
Query Match
Best Local Similarity 53.5%; Score 53; DB 11; Length 143;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 INPRNDYTEXNONFK 16
DB 51 IDPSDYTYNOKFK 65
RESULT 7
O815A3 PRELIMINARY; PRT; 1824 AA.
AC 0815A3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Asparagine-rich protein, putative.
GN PFI1530W.
OS Plasmodium falciparum (Isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=12368864;
Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguolli S.,
RA Perlea M., Allen T., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M.A., Faloutsos L.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum." ;
RL Nature 419:498-511(2002).
DR EMBL; AE014849; AAN36392.1; -.
SQ SEQUENCE 1824 AA; 212658 MW; 6ECED0EA6960084E CRC64;
Query Match
Best Local Similarity 53.0%; Score 52.5; DB 5; Length 1824;
Matches 10; Conservative 4; Mismatches 2; Indels 9; Gaps 1;
QY 2 INPR-----NDYTEXNONFKD 17
DB 1751 LNDKFNLDLNNNTYNTDNDNFKD 1775
RESULT 8
Q924Q2 PRELIMINARY; PRT; 142 AA.
AC 0924Q2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE V303-D-J-C mu protein (Fragment).
GN V303-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP)." ;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069910; BAB63926.1; -.
DR InterPro; IPR007110; I9-1like.
DR InterPro; IPR003006; I9_MHC.
DR InterPro; IPR003596; I9_V.
DR Pfam; PF00047; I9; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1 142
SQ SEQUENCE 142 AA; 15684 MW; 04880F9FC0FBD7B CRC64;
Query Match
Best Local Similarity 52.5%; Score 52; DB 11; Length 142;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 INPRNDYTEXNONFK 16
DB 51 IDPSDIYTDYNOEFK 65
RESULT 9
O8Y7Z9 PRELIMINARY; PRT; 138 AA.
AC 08Y7Z9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein IMO1120.
GN IMO1120.

```
OS Listeria monocytogenes.
OC Bacteria: Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnock C., Amend A.,
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Darivar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauguet O.,
RA Entian K.-D., Faehl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hahn T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Medjert H.,
RA Nodtlex G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT *Comparative genomics of Listeria species.*;
RL Science 294:849-852(2001);
DR EMBL: AL591977; CAC99198.1; -
DR Listlist: LMO01120; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 138 AA; 15807 MW; DC1AC34EE77A7AC2 CRC64;

Query Match 51.5%; Score 51; DB 16; Length 138;
Best Local Similarity 52.9%; Pred. No. 3.1;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNONFKD 17
DB 66 YINPKOFWEFENOPFLD 82
||||: |||||

RESULT 10
Q99L25 PRELIMINARY; PRT; 473 AA.
ID Q99L25;
AC Q99L25;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 181006009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR HSP; P01842; FAB.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00406; Igv; 1.
DR PROSITE: PS50835; Ig_LIKE; 4.
DR PROSITE: PS00290; Ig_MHC; 1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match 51.5%; Score 51; DB 11; Length 473;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YINPRNDYTEYNONFK 16
DB 69 YIIPRDSCTKYNFK 84
||||: |||||

RESULT 11
Q91WT3 PRELIMINARY; PRT; 481 AA.
ID Q91WT3
```

```
AC Q91WT3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 52.0 kDa protein.
GN IGH-VJ558 OR A1893585.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC013488; AAH13488.1; -
DR MGD; MGI:96486; Igh-VJ558.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00406; Igv; 1.
DR PROSITE: PS50835; Ig_LIKE; 4.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52022 MW; 4EEB5C253038B718 CRC64;

Query Match 51.5%; Score 51; DB 11; Length 481;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 INPRNDYTEYNONFK 16
DB 70 IDPFDSYSTYNOKFK 84
||||: |||||

RESULT 12
Q91WR1 PRELIMINARY; PRT; 488 AA.
ID Q91WR1;
AC Q91WR1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 53.0 kDa protein.
GN IGH-VJ558 OR A1893585.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strauberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC013539; AAH13539.1; -
DR MGD; MGI:96486; Igh-VJ558.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00406; Igv; 1.
DR PROSITE: PS50835; Ig_LIKE; 4.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 52964 MW; F12068460B4009D CRC64;

Query Match 51.5%; Score 51; DB 11; Length 488;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 INPRNDYTEYNONFK 16
DB 70 INPNGSTSYNOKFK 84
||||: |||||
```

```
RESULT 13
ID 050330 PRELIMINARY; PRT; 823 AA.
AC 050330: 093732;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Trk protein.
GN Trk.
OS Escherichia coli.
OC plasmid R388.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCB1_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K-12; PLASMD=R388.
RX MEDLINE=97467347; PubMed=9325277;
RA Bolland S., Lloza M., de la Cruz F.;
RT "Genetic organization of the region involved in conjugative pilus
RT synthesis, export and assembly of the Incw plasmid R388.";
RL J Biol. Chem. 272:25583-25590(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K-12; PLASMD=R388;
RA de la Cruz F.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC PLASMD=R388;
RA Rabel C., Lanka E.;
RT "Direct Submision.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; X81123; CA57024.2; -
DR EMBL; AJ316617; CAC78962.1; -
DR InterPro; IPR004346; CAGE_TrbE_VtrB.
DR Pfam; PF03135; CAGE_TrbE_VtrB.1.
DR TIGRfams; TIGR00929; VtrB4_Cage; 1.
KW Plasmid.
SQ SEQUENCE 823 AA; 93660 MW; 5AE3115FECFA8208 CRC64;

Query Match 50.5%; Score 50; DB 2; Length 823;
Best Local Similarity 57.1%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 NPNRDYTYEYNQFK 16
DB 714 NPKADYEDYTGFK 727

RESULT 14
ID 022624 PRELIMINARY; PRT; 1170 AA.
AC 022624;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE T21B10.3 protein.
GN T21B10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Baynes C.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9651916;
RA none;
```

```
RT "Genome sequence of the nematode C.elegans: A platform for
RT Investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z68318; CA92693.2; -
DR WormPep; T21B10.3; CE23993
SQ SEQUENCE 1170 AA; 135533 MW; 461F6BA1A2B6BF04 CRC64;

Query Match 50.5%; Score 50; DB 5; Length 1170;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PRNDYTYEQ 13
DB 984 PRNDYSEYDQ 993

RESULT 15
ID 0924P8 PRELIMINARY; PRT; 140 AA.
AC 0924P8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE V23-D-J-C mu protein (Fragment).
GN V23-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069917; BAB63933.1; -
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003006; IG-MHC.
DR InterPro; IPR003596; IG-V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 140 AA; 15392 MW; 904C80C82548C936 CRC64;

Query Match 49.5%; Score 49; DB 11; Length 140;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 INPRNDYTYEYNQFK 16
DB 51 INPNSGTNYMEKFK 65

Search completed: October 7, 2003, 19:20:03
Job time : 18.9414 secs
```


[illegible]

XX	WP1: 1996-139454/14.
DR	N-PSDB; AART15804.
XX	
PT	Chimeric and humanised LL2 antibodies - used to produce conjugates
PT	for the therapy and diagnosis of B-cell lymphoma(s) and
PT	leukaemia(s).
XX	
PS	Claim 5; Page 39; 70pp; English.
XX	
CC	The complementarity determining regions (CDRs) of mouse monoclonal
CC	antibody (MAb) LL2 VK (AAR92215) and VH (AAR92216) regions were
CC	recombinantly linked to the framework sequences of human VK and VH
CC	regions, respectively, to give humanised LL2 VK (AAR92217) and VH
CC	(AAR92218). These were subsequently linked, respectively, to human
CC	kappa and Igcl constant regions. A humanised MAb was obtd. that
CC	retained the B-lymphoma and leukaemia cell targeting and
CC	internalisation characteristics of the parental LL2 MAb, and which
CC	exhibited a lowered HAMA reaction. It can be linked to e.g. a
CC	cytostatic agent for therapeutic appln.
XX	
SO	Sequence 116 AA;
Query Match	100.0%; Score 38; DB 17; Length 116;
Best Local Similarity	100.0%; Pred. No. 2,3;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 RDITTFY 7
Db	99 RDITTFY 105
RESULT 3	
AAR92219	
ID	AAR92219 standard; Protein; 116 AA.
XX	
AC	AAR92219;
XX	
DT	28-MAY-1996 (first entry)
XX	
DE	Humanised LL2 MAb VH region (version hLL2-1).
XX	
KW	Humanised antibody; monoclonal antibody; MAb; LL2; B-cell lymphoma;
KW	leukaemia; therapy; diagnosis; complementarity determining region;
KW	CDR; antibody engineering.
XX	
OS	Synthetic.
XX	
FH	Key
FH	Region
FT	Location/Qualifiers
FT	31..35
FT	/label= CDR1
FT	50..66
FT	/label= CDR2
FT	99..105
FT	/label= CDR3
XX	
PN	WO9604925-A1.
XX	
PD	22-FEB-1996.
XX	
PF	11-AUG-1995; 95WO-US09641.
XX	
PR	12-AUG-1994; 94US-0289576.
XX	
PA	(IMMU-) IMMUNOMEDICS INC.
XX	
PI	Hansen H, Leung S;
XX	
DR	WP1: 1996-139454/14.
XX	
PT	Chimeric and humanised LL2 antibodies - used to produce conjugates
PT	for the therapy and diagnosis of B-cell lymphoma(s) and
PT	leukaemia(s).

```

XX PS Example 1; Page 40; 70pp; English.
XX CC The complementarily determining regions (CDRs) of mouse monoclonal
XX CC antibody (Mab) LL2 VK (AAR92215) and VH (AAR92216) regions were
XX CC recombinantly linked to the framework sequences of human VK and VH
XX CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
XX CC (AAR92218). In an alternative version, a glutamine was introduced
XX CC at position 5 of the humanised VH (AAR92219) to include a PstI site
XX CC useful for subcloning. The humanised VK and VH were subsequently
XX CC linked, respectively, to human kappa and IgG1 constant regions. A
XX CC humanised Mab was obtd. that retained the B-lymphoma and leukaemia
XX CC cell targeting and internalisation characteristics of the parental
XX CC LL2 Mab, and which exhibited a lowered HAMA reaction. It can be
XX CC linked to a cytostatic agent for therapeutic appln.
XX SQ Sequence 116 AA;

Query Match 100.0%; Score 38; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
Db 99 RDITTFY 105

RESULT 4
AAW27696
ID AAW27696 standard; Protein; 116 AA.
AC AAW27696;
XX 14-APR-1998 (first entry)
XX DE Variable heavy chain of Mab LL2.
XX DE
XX KW Variable heavy chain; B cell; monoclonal antibody; Mab; LL2;
XX KW B cell lymphoma; lymphocytic leukaemia cell; murine;
XX KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
XX KW chronic lymphocytic leukaemia.
XX OS Mus sp.
XX OS Synthetic.
XX FH Key
XX FH Location/Qualifiers
XX FT 31..35
XX FT /note= "complementarity determining region 1"
XX FT 50..66
XX FT /note= "complementarity determining region 2"
XX FT 99..105
XX FT /note= "complementarity determining region 3"
XX PN W09734632-A1.
XX PD 25-SEP-1997.
XX PF 19-MAR-1997; 97WO-US04196.
XX PR 20-MAR-1996; 96US-0013709.
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PI Hansen H, Leung S, Qu Z;
XX DR WPI: 1997-479995/44.
XX DR N-PSDB: AAT88129.
XX PT Monoclonal antibody engineered to contain glycosylation site - in
XX PT non-Fc constant heavy or light chain region, useful to diagnose or
XX PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX PS Example 3; Fig 4B; 88pp; English.

```

```

XX CC The present sequence is the variable heavy chain of the
XX CC B cell specific monoclonal antibody (Mab) LL2, which contains an
XX CC engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly
XX CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
XX CC murine Mab. The Mab can be used to diagnose or treat B
XX CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
XX CC lymphocytic leukaemia. The glycosylation site allows a label or
XX CC therapeutic agent of increased size to be conjugated to the
XX CC carbohydrate moiety, without affecting the Mab's binding affinity
XX CC or specificity.
XX SQ Sequence 116 AA;

Query Match 100.0%; Score 38; DB 18; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
Db 99 RDITTFY 105

RESULT 5
AAW27698
ID AAW27698 standard; Protein; 116 AA.
AC AAW27698;
XX 14-APR-1998 (first entry)
XX DE Variable heavy chain of Mab hLL2.
XX DE
XX KW Variable heavy chain; B cell; monoclonal antibody; Mab; hLL2;
XX KW B cell lymphoma; lymphocytic leukaemia cell; murine; humanised;
XX KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
XX KW chronic lymphocytic leukaemia.
XX OS Chimeric - Mus sp.
XX OS Chimeric - Homo sapiens.
XX FH Key
XX FH Location/Qualifiers
XX FT 31..35
XX FT /note= "complementarity determining region 1"
XX FT 50..66
XX FT /note= "complementarity determining region 2"
XX FT 99..105
XX FT /note= "complementarity determining region 3"
XX PN W09734632-A1.
XX PD 25-SEP-1997.
XX PF 19-MAR-1997; 97WO-US04196.
XX PR 20-MAR-1996; 96US-0013709.
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PI Hansen H, Leung S, Qu Z;
XX DR WPI: 1997-479995/44.
XX DR N-PSDB: AAT88131.
XX PT Monoclonal antibody engineered to contain glycosylation site - in
XX PT non-Fc constant heavy or light chain region, useful to diagnose or
XX PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX PS Example 3; Fig 5B; 88pp; English.
XX CC The present sequence is the variable heavy chain of the
XX CC B cell specific monoclonal antibody (Mab) hLL2. hLL2 is a highly
XX CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell

```

CC humanised murine Mab. The Mab can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia.

SO Sequence 116 AA;

Query Match 100.0%; Score 38; DB 18; Length 116;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7

Db 99 RDITTFY 105

RESULT 6
AAU30864
ID AAU30864 standard; Protein; 146 AA.

AC AAU30864;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #1355.

KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN MO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001MO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -

PS Claim 20; Page 363; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

SO Sequence 146 AA;

Query Match 86.8%; Score 33; DB 22; Length 146;
Best Local Similarity 71.4%; Pred. No. 34;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7

Db 28 RDITTFY 34

RESULT 7
AAB94060
ID AAB94060 standard; Protein; 203 AA.

AC AAB94060;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:14236.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 8; SEQ ID 14236; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

SO Sequence 203 AA;

Query Match 86.8%; Score 33; DB 22; Length 203;

Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
|||:
Db 85 RDVTTF 91

RESULT 8
AAG07543
ID AAG07543 standard; Protein: 277 AA.
XX
AC AAG07543;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4737.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145148.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149729.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160880.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161361.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match Best Local Similarity 86.8%; Score 33; DB 21; length 277;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDIITTFY 7
|::|||
Db 142 REVTTTFY 148

RESULT 9
AAC53675
ID AAC53675 standard; Protein: 277 AA.
XX
AC AAC53675;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 68358.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135353.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.8%; Score 33; DB 21; Length 277;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 RDIITFFY 7
Db 142 REVITFFY 148

RESULT 10
AAG07542
ID AAG07542 standard; Protein; 333 AA.
AC
XX AAG07542;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 4736.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155639.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160817.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.8%; Score 33; DB 21; Length 333;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
|:|||||
Db 198 REVTFY 204

RESULT 11
AAG53674
ID AAG53674 standard; Protein: 333 AA.

AC AAG53674;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 68357.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.
OS
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159884.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.8%; Score 33; DB 21; Length 333;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDTTFY 7
Db 198 REVTFY 204

RESULT 12
AAAG99332 standard; Protein; 337 AA.
ID AAAG99332;

AC AAAG99332;
XX XX

DT 28-SEP-2001 (first entry)
XX XX

DE Maize patatin homolog corn5 pep.
XX XX

KW Insecticidal; immunosuppressive; potato; patatin; enzyme;
KM allergenicity; larva growth; lipid acyl hydrolase; insecticide.
XX XX

OS Zea mays.
XX XX

PN WO200149834-A2.
XX XX

PD 12-JUL-2001.
XX XX

PF 05-JAN-2001; 2001WO-US00342.

XX 06-JAN-2000; 2000US-0174669.
PR (MONS) MONSANTO CO.
XX
XX Alldhal MF, Astwood JD, McWherter CA, Sampson HA;
PI WPI; 2001-441874/47.
XX
XX Modified potato patatin proteins with reduced antigenicity, useful as
PT insecticides for controlling e.g. round worm and root worm -
XX
XX Claim 55; Pages 222-223; 223pp; English.
XX
XX The present invention relates to modified potato patatins that maintain
CC enzymatic and insecticidal activity but which have reduced allergenicity.
CC Groups (especially Tyr) which bind to anti-patatin antibodies were
CC identified and glycosylation sites involved in antibody binding were
CC removed via site directed mutagenesis. The patatins stunt the growth of
CC larvae so that maturation is prevented or delayed. The patatins also have
CC non-specific lipid acyl hydrolase activity. The modified patatins are
CC also useful for inhibiting the activity of corn round worms.
CC Deallergenised protein can be used as insecticides, as nutritional
CC supplements and as immunising agents. The present sequence was used to
CC illustrate the present invention.
XX
SQ Sequence 337 AA;

Query Match 86.8%; Score 33; DB 22; Length 337;
Best Local Similarity 71.4%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
Db 95 KDLTFY 101

RESULT 13

AAB93724 ID AAB93724 standard; Protein; 343 AA.

XX AAB93724;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13343.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 13343; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 343 AA;

Query Match 86.8%; Score 33; DB 22; Length 343;
Best Local Similarity 71.4%; Pred. No. 86;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
Db 225 RDTTFY 231

RESULT 14

AAG53673 ID AAG53673 standard; Protein; 361 AA.

XX AAG53673;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 68356.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0128845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.

PR 20-JUL-1999: 99US-0144684.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 10-AUG-1999: 99US-0147935.
PR 11-AUG-1999: 99US-0148171.
PR 12-AUG-1999: 99US-0148319.
PR 13-AUG-1999: 99US-0148361.
PR 13-AUG-1999: 99US-0148365.
PR 16-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.

PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match

Best Local Similarity 86.8%; Score 33; DB 21; Length 361;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDIITFFY 7
|::|||
DB 226 REVTFY 232

RESULT 15

AA607541
ID AA607541 standard; Protein; 362 AA.

AC AA607541;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4735.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-012348.

PR 23-MAR-1999; 990S-0125788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.

PR 23-APR-1999; 990S-0130891.

PR 28-APR-1999; 990S-0131449.

PR 30-APR-1999; 990S-0132048.

PR 30-APR-1999; 990S-0132407.

PR 04-MAY-1999; 990S-0132484.

PR 05-MAY-1999; 990S-0132485.

PR 06-MAY-1999; 990S-0132486.

PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134321.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142803.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143342.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147453.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.8%; Score 33; DB 21; Length 362;
Best Local Similarity 71.4%; Pred. No. 91;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RDITTFY 7
Db 227 REVTTTFY 233

Search completed: October 7, 2003, 19:13:59
Job time : 10.8103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:41 : Search time 3.62069 Seconds
(without alignments)
305.878 Million cell updates/sec

Title: US-09-988-013a-4_COPY_99_105
Perfect score: 38
Sequence: 1 RDITTFY 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgnt2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgnt2_6/ptodata/1/pubppaa/PCF_NEW_PUB.pep:*
- 3: /cgnt2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgnt2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgnt2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgnt2_6/ptodata/1/pubppaa/PCFUS_PUBCOMB.pep:*
- 7: /cgnt2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgnt2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgnt2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgnt2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgnt2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgnt2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgnt2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgnt2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgnt2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgnt2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgnt2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgnt2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	38	100.0	116	10	US-09-741-843-4
2	38	100.0	116	10	US-09-741-843-8
3	38	100.0	116	10	US-09-741-843-9
4	38	100.0	116	11	US-09-894-839-4
5	38	100.0	116	11	US-09-894-839-8
6	38	100.0	116	11	US-09-894-839-21
7	38	100.0	116	11	US-09-888-013a-4
8	38	100.0	116	11	US-09-888-013a-8
9	38	100.0	116	11	US-09-988-013a-9
10	38	86.8	2644	12	US-10-300-453a-39
11	32	84.2	53	10	US-09-796-692-1784
12	32	84.2	53	10	US-09-796-692-2530
13	32	84.2	53	15	US-10-040-862-1784
14	32	84.2	53	15	US-10-040-862-2530
15	32	84.2	218	9	US-09-925-301-1242

16	32	84.2	406	12	US-10-334-038-16	Sequence 16, Appl
17	32	84.2	406	15	US-10-228-897-12	Sequence 12, Appl
18	31	81.6	543	15	US-10-156-761-8387	Sequence 8387, Ap
19	31	81.6	565	12	US-10-132-134-26	Sequence 26, Appl
20	31	81.6	1395	10	US-09-808-602-67	Sequence 67, Appl
21	31	81.6	1395	11	US-09-800-198-56	Sequence 56, Appl
22	31	81.6	1395	12	US-10-289-776-15	Sequence 15, Appl
23	30	78.9	544	9	US-09-809-342a-2	Sequence 2, Appl
24	30	78.9	544	11	US-09-946-374-118	Sequence 118, App
25	30	78.9	544	12	US-10-015-387a-118	Sequence 118, App
26	30	78.9	544	12	US-10-006-130a-118	Sequence 118, App
27	30	78.9	544	12	US-10-199-672-274	Sequence 274, App
28	30	78.9	544	12	US-10-006-172a-118	Sequence 118, App
29	30	78.9	544	12	US-10-187-749-274	Sequence 274, App
30	30	78.9	544	12	US-10-194-457-274	Sequence 274, App
31	30	78.9	544	12	US-10-184-642-274	Sequence 274, App
32	30	78.9	544	12	US-10-196-747-274	Sequence 274, App
33	30	78.9	544	12	US-10-015-392a-118	Sequence 274, App
34	30	78.9	544	12	US-10-017-353a-118	Sequence 118, App
35	30	78.9	544	12	US-10-173-689-274	Sequence 118, App
36	30	78.9	544	12	US-10-173-690-274	Sequence 274, App
37	30	78.9	544	12	US-10-173-691-274	Sequence 274, App
38	30	78.9	544	12	US-10-173-692-274	Sequence 274, App
39	30	78.9	544	12	US-10-173-694-274	Sequence 274, App
40	30	78.9	544	12	US-10-173-698-274	Sequence 274, App
41	30	78.9	544	12	US-10-173-699-274	Sequence 274, App
42	30	78.9	544	12	US-10-173-707-274	Sequence 274, App
43	30	78.9	544	12	US-10-174-569-274	Sequence 274, App
44	30	78.9	544	12	US-10-174-583-274	Sequence 274, App
45	30	78.9	544	12	US-10-174-587-274	Sequence 274, App

ALIGNMENTS

RESULT 1
US-09-741-843-4
; Sequence 4, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-con
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-741-843-4
Query Match 100.0%; Score 38; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RDITTFY 7
Db 99 RDITTFY 105
RESULT 2
US-09-741-843-8
; Sequence 8, Application US/09741843

```
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LY
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-8
```

```
Query Match          100.0%; Score 38; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 RDITTFY 7
        |||||
Db       99 RDITTFY 105
```

```
RESULT 3
US-09-741-843-9
; Sequence 9, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LY
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-741-843-9
```

```
Query Match          100.0%; Score 38; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 RDITTFY 7
        |||||
Db       99 RDITTFY 105
```

```
RESULT 4
US-09-894-839-4
; Sequence 4, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
```

```
; APPLICANT: HANSEN, Hans
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-894-839-4
```

```
Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 RDITTFY 7
        |||||
Db       99 RDITTFY 105
```

```
RESULT 5
US-09-894-839-8
; Sequence 8, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-8
```

```
Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 RDITTFY 7
        |||||
Db       99 RDITTFY 105
```

```
RESULT 6
US-09-894-839-21
; Sequence 21, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
```

```
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: US 60/013,709
;; PRIOR FILING DATE: 1996-03-20
;; NUMBER OF SEQ ID NOS: 47
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 21
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-894-839-21

Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RDITTFY 7
        |||||
Db      99 RDITTFY 105

RESULT 7
US-09-988-013A-4
; Sequence 4, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-988-013A-4

Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RDITTFY 7
        |||||
Db      99 RDITTFY 105

RESULT 8
US-09-988-013A-8
; Sequence 8, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
```

```
;; PRIOR APPLICATION NUMBER: US 08/690,102
;; PRIOR FILING DATE: 1996-07-06
;; PRIOR APPLICATION NUMBER: US 08/289,576
;; PRIOR FILING DATE: 1994-08-12
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8
;; LENGTH: 116
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-988-013A-8

Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RDITTFY 7
        |||||
Db      99 RDITTFY 105

RESULT 9
US-09-988-013A-9
; Sequence 9, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-988-013A-9

Query Match          100.0%; Score 38; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RDITTFY 7
        |||||
Db      99 RDITTFY 105

RESULT 10
US-10-300-453A-39
; Sequence 39, Application US/10300453A
; Publication No. US20030165934A1
; GENERAL INFORMATION:
; APPLICANT: ELEDGE, STEPHEN J.
; APPLICANT: CORTEZ, DAVID K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN CHECKPOINT SIGNALING
; FILE REFERENCE: P023390US
; CURRENT APPLICATION NUMBER: US/10/300,453A
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 60/331,821
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 47
```

```
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 39
;; LENGTH: 2644
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-300-453A-39

Query Match      86.8%; Score 33; DB 12; Length 2644;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RDITTFY 7
      |||||
Db      1814 RDITTFY 1820

RESULT 11
US-09-796-692-1784
; Sequence 1784, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1784
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1784

Query Match      84.2%; Score 32; DB 10; Length 53;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RDITTFY 7
      |||||
Db      32 RDITTFY 38

RESULT 12
US-09-796-692-2530
; Sequence 2530, Application US/09796692
```

```
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2530
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-796-692-2530

Query Match      84.2%; Score 32; DB 10; Length 53;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RDITTFY 7
      |||||
Db      32 RDITTFY 38

RESULT 13
US-10-040-862-1784
; Sequence 1784, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
```

```
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1784
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-040-862-1784
```

```
Query Match      84.2%; Score 32; DB 15; Length 53;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 RDITFFY 7
      ||| |||
Db      32 RDITFFY 38
```

```
RESULT 14
US-10-040-862-2530
; Sequence 2530, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
```

```
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2530
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-040-862-2530
```

```
Query Match      84.2%; Score 32; DB 15; Length 53;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 RDITFFY 7
      ||| |||
Db      32 RDITFFY 38
```

```
RESULT 15
US-09-925-301-1242
; Sequence 1242, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1242
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1242
```

```
Query Match      84.2%; Score 32; DB 9; Length 218;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 RDITFFY 7
      ||| |||
Db      197 RDITFFY 203
```

Search completed: October 7, 2003, 19:24:28
Job time : 4.62069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:07:11 ; Search time 2.6069 Seconds
(without alignments)
258.231 Million cell updates/sec

Title: US-09-988-013a-4_COPY_99_105

Perfect score: 38

Sequence: 1 RDITTFY 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	89.5	98	D89856	hypothetical prote
2	33	86.8	168	T46298	hypothetical prote
3	33	86.8	333	D85068	D123-like protein
4	32	84.2	377	VHWH79	nucleocapsid prote
5	32	84.2	398	S13269	translation initia
6	32	84.2	406	F1MS4A	translation initia
7	32	84.2	406	S33681	translation initia
8	32	84.2	407	S00985	translation initia
9	31	81.6	35	H95100	hypothetical prote
10	31	81.6	166	T50172	hypothetical prote
11	31	81.6	278	A42400	probable transcrip
12	31	81.6	279	E69226	hypothetical prote
13	31	81.6	307	G69501	UDP-glucose 4-epim
14	31	81.6	447	S44133	histidine kinase -
15	31	81.6	695	E87397	transglycosylase,
16	31	81.6	830	T04848	protein kinase hom
17	31	81.6	6359	T31679	bacteracl synthe
18	30	78.9	107	B48677	Ig kappa chain V-J
19	30	78.9	166	S52308	pall protein - pse
20	30	78.9	167	S58217	outer membrane pro
21	30	78.9	168	G83525	outer membrane pro
22	30	78.9	1549	D86338	protein F5M15.18 l
23	29	76.3	203	A36365	transforming prote
24	29	76.3	276	D75015	hypothetical prote
25	29	76.3	285	T15252	hypothetical prote
26	29	76.3	299	T32094	hypothetical prote
27	29	76.3	302	B35271	protocatalase 4,
28	29	76.3	341	C86246	hypothetical prote
29	29	76.3	357	ATQO	actin - Oxytricha

30	29	76.3	375	2	A31134	actin, macronuclea
31	29	76.3	399	2	T01387	oxidoreductase hom
32	29	76.3	428	2	A89950	glutamate-1-semal
33	29	76.3	433	2	B90419	hypothetical prote
34	29	76.3	441	2	T50436	hypothetical prote
35	29	76.3	510	2	E71695	hypothetical prote
36	29	76.3	523	2	AD0506	probable secreted
37	29	76.3	588	2	D90750	hypothetical prote
38	29	76.3	588	2	B85614	hypothetical prote
39	29	76.3	588	2	AC0611	transport ATP-bind
40	29	76.3	588	2	F64827	ABC-type transport
41	29	76.3	594	2	AC0167	transport ATP-bind
42	29	76.3	642	2	T28866	hypothetical prote
43	29	76.3	716	2	C53292	penicillin-binding
44	29	76.3	880	2	D69427	conserved hypothet
45	29	76.3	971	2	S50912	probable membrane

ALIGNMENTS

RESULT 1

D89856
hypothetical protein SAS023 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89856
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Tl, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <KUR>
A:Cross-references: GB:BA000018; PID:913700710; PIDN:BA042007.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SAS023

Query Match 89.5%; Score 34; DB 2; Length 98;
Best Local Similarity 71.4%; Pred. No. 3.2;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
Db 8 KDVTTFY 14

RESULT 2

T46298
hypothetical protein DKFZp4341110.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46298
R:Duisterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23035
A:Accession: T46298
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-168 <AAA>
A:Cross-references: EMBL:AL137744
A:Experimental source: adult testis; clone DKFZp43411110
C:Genetics:
A:Note: DKFZp4341110.1

Query Match 86.8%; Score 33; DB 2; Length 168;
Best Local Similarity 71.4%; Pred. No. 9.1;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
DB 50 RDVTFY 56

RESULT 3

D85068
D123-like protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C/Accession: D85068
R/anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: D85068
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-333 <STO>
A/Cross-references: GB:NC_001268; NID:97267304; PIDN:CA81086.1; GSPDB:GN00140
C/Genetics:
A/Gene: AT4G05440
A/Map position: 4
C/Superfamily: Saccharomyces cerevisiae hypothetical protein YLR215c

Query Match 86.8%; Score 33; DB 2; Length 333;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
DB 198 RDVTFY 204

RESULT 4

VH1H79
nucleocapsid protein - feline infectious peritonitis virus (strain 79-1146)
C/Species: feline infectious peritonitis virus
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: B38498
R/Vennema, H.; De Groot, R.J.; Harbour, D.A.; Horzinek, M.C.; Spaan, W.J.M.
Virology 181, 327-335, 1991
A/Title: Primary structure of the membrane and nucleocapsid protein genes of feline infe
A/Reference number: A38498; MUID:91134997; PMID:1847259
A/Accession: B38498
A/Molecule type: genomic RNA
A/Residues: 1-377 <VEN>
A/Cross-references: EMBL:X56496; NID:958918; PIDN:CAA39851.1; PID:958920
C/Superfamily: coronavirus nucleocapsid protein
C/Keywords: glycoprotein; nucleocapsid
F:134,154,158,172/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.2%; Score 32; DB 1; Length 377;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DITTFY 7
DB 241 DVTTFY 246

RESULT 5

S13269
translation initiation factor eIF-4A.I - rabbit (fragment)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 19-Mar-1997 #sequence_revision 17-Apr-1998 #text_change 02-Feb-2001
C/Accession: S13269; S78482
R/Conroy, S.C.; Dever, T.E.; Owens, C.L.; Merrick, W.C.
Arch. Biochem. Biophys. 282, 363-371, 1990
A/Title: Characterization of the 46,000-dalton subunit of eIF-4F.
A/Reference number: S13269; MUID:91053155; PMID:2241157
A/Accession: S13269
A/Molecule type: mRNA

A/Residues: 5-394 <CON1>
A/Experimental source: reticulocyte
A/Accession: S78482
A/Molecule type: protein
A/Residues: 1-25;38-53;61-102;142-170;183-194;219-225;275-283;309-345;361-398 <CON2>
A/Experimental source: reticulocyte
C/Comment: This is one of the many factors involved in the complex assembly process o
a protein complex involved in "cap" recognition.
C/Genetics:
A/Gene: eIF-4A1
C/Function:

A/Description: translation initiation
C/Superfamily: translation initiation factor eIF-4A
C/Keywords: ATP; blocked amino end; nucleotide binding; P-loop; protein biosynthesis;
F:1-398/Product: translation initiation factor eIF-4A.I (fragment) #status experiment
F:68-75/Region: nucleotide-binding motif A (P-loop)
F:170-175/Region: nucleotide-binding motif B
F:174-177/Region: DEAD motif
F:74/Binding site: ATP (Lys) #status predicted

Query Match 84.2%; Score 32; DB 2; Length 398;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RDITTFY 7
DB 377 RDITTFY 383

RESULT 6

FIMS4A
translation initiation factor eIF-4A I - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1988 #sequence_revision 09-Apr-1998 #text_change 19-Jan-2001
C/Accession: J50039; A24267; B24267; S00986; A58767
R/Reddy, N.S.; Roth, W.W.; Bragg, P.W.; Wabba, A.J.
Gene 70, 231-243, 1988
A/Title: Isolation and mapping of a gene for protein synthesis initiation factor 4A a
A/Reference number: J50039; MUID:89108007; PMID:3215517
A/Accession: J50039
A/Molecule type: DNA
A/Residues: 1-299, 'C', '301-406 <RED>
A/Cross-references: GB:L13611; NID:9556306; PIDN:AA50407.1; PID:9556308
A/Note: the authors translated the codon GAG for residue 143 as Gly
R:Nielsen, P.J.; McMaster, G.K.; Trachsel, H.
Nucleic Acids Res. 13, 6867-6880, 1985
A/Title: Cloning of eukaryotic protein synthesis initiation factor genes: Isolation a
A/Reference number: A24267; MUID:86041884; PMID:3840589
A/Accession: A24267
A/Molecule type: mRNA
A/Residues: 17-406 <NIE>
A/Cross-references: GB:X03039; NID:950812; PIDN:CAA26842.1; PID:950814; GB:X03040; NI
A/Accession: B24267
A/Molecule type: mRNA
A/Residues: 37-406 <NIE2>
A/Cross-references: GB:X03039; NID:950812; PIDN:CAA26843.1; PID:950815; GB:X03040; NI
R:Nielsen, P.J.; Trachsel, H.
EMBO J. 7, 2097-2105, 1988
A/Title: The mouse protein synthesis initiation factor 4A gene family includes two re

A/Reference number: S00985; MUID:88328998; PMID:3046931
A/Accession: S00986
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 'W', '1-154', 'VI', '157-406 <NIE3>
A/Cross-references: EMBL:X14421; NID:950810
A/Accession: A58767
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-8 <NIE4>
A/Cross-references: EMBL:X14421; NID:950810; PIDN:CAA32584.1; PID:9581984
C/Comment: This is one of the many factors involved in the complex assembly process o
a protein complex involved in "cap" recognition.

C:Genetics:
A:Gene: MGI:Elf4a1
A:Cross-references: MGI:95303
A:Map position: 11:39.0
A:Introns: 8/2; 24/3; 68/3; 115/3; 172/1; 208/3; 256/3; 302/3; 332/3; 359/2
C:Superfamily: translation initiation factor eIF-4A
C:Keywords: ATP; nucleotide binding; P-loop; protein biosynthesis; RNA binding
F:76-83/Region: nucleotide-binding motif A (P-loop)
F:178-183/Region: nucleotide-binding motif B
F:182-185/Region: DEAD motif
F:82/Binding site: ATP (Lys) #status predicted

Query Match 84.2%; Score 32; DB 1; Length 406;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
DB 385 RDITTFY 391

RESULT 7
S33681
translation initiation factor eIF-4A.I - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 17-Nov-1995 #text_change 02-Feb-2001
C:Accession: S33681
R:Kim, N.S.; Kato, T.; Abe, N.; Kato, S.
Nucleic Acids Res. 21, 2012, 1993
A:Title: Nucleotide sequence of human cDNA encoding eukaryotic initiation factor 4A1.
A:Reference number: S33681; MUID:93261841; PMID:8493113
A:Accession: S33681
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-406 <KMT>
A:Cross-references: EMBL:DJ3748; MUD:9219402; PTDN:BA02897.1; PTD:9219403
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
C:Superfamily: translation initiation factor eIF-4A
C:Keywords: ATP; nucleotide binding; P-loop; protein biosynthesis; RNA binding
F:76-83/Region: nucleotide-binding motif A (P-loop)
F:178-183/Region: nucleotide-binding motif B
F:182-185/Region: DEAD motif

Query Match 84.2%; Score 32; DB 2; Length 406;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
DB 385 RDITTFY 391

RESULT 8
S00985
translation initiation factor eIF-4A II - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 02-Feb-2001
C:Accession: S00985; S12470
R:Nielsen, P.J.; Trachsel, H.
EMBO J. 7, 2097-2105, 1988
A:Title: The mouse protein synthesis initiation factor 4A gene family includes two related
A:Reference number: S00985; MUID:8832898; PMID:3046931
A:Accession: S00985
A:Molecule type: DNA
A:Residues: 1-407 <NTE>
A:Cross-references: EMBL:X12507
R:Nielsen, P.J.
Submitted to the EMBL Data Library, July 1989
A:Reference number: S12470
A:Accession: S12470
A:Molecule type: DNA
A:Residues: 1-199, 'RV', 202-293, 'Q', 295, 'TV', 298-386, 'H', 388-402, 'G', 404-407 <NIE2>
A:Cross-references: EMBL:X12507; MUD:950822; PTDN:CAA31025.1; PID:950823

C:Superfamily: translation initiation factor eIF-4A
C:Keywords: ATP; nucleotide binding; P-loop; protein biosynthesis; RNA binding
F:77-84/Region: nucleotide-binding motif A (P-loop)
F:179-184/Region: nucleotide-binding motif B
F:183-186/Region: DEAD motif
F:83/Binding site: ATP (Lys) #status predicted

Query Match 84.2%; Score 32; DB 2; Length 407;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
DB 386 RDITTFY 392

RESULT 9
H95100
hypothetical protein SP0874 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95100
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, U.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <KUR>
A:Cross-references: GB:AE005672; PTDN:AAK75001.1; PTD:914972347; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0874

Query Match 81.6%; Score 31; DB 2; Length 35;
Best Local Similarity 83.3%; Pred. No. 5;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DITTFY 7
DB 21 DITTFY 26

RESULT 10
T50172
hypothetical protein SPAC227.16c [imported] - fission yeast (Schizosaccharomyces pombe
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50172
R:Zimmermann, W.; Wandut, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, November 1999
A:Reference number: Z25036
A:Accession: T50172
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-166 <ZIK>
A:Cross-references: EMBL:AL133156; PTDN:CAB61465.1; GSPDB:GN00066; SPDB:SPAC227.16c
A:Experimental source: strain 972h(-); cosmid c227
C:Genetics:
A:Gene: SPDB:SPAC227.16c
A:Map position: 1
A:Introns: 33/3; 53/2; 98/2; 147/1

Query Match 81.6%; Score 31; DB 2; Length 166;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
DB 21 DITTFY 26

Db 82 RDITTY 88

RESULT 11

A42400

probable transcription regulator msmr [similarity] - Streptococcus mutans

N.Alternate names: multiple sugar metabolism protein msmr

C:Species: Streptococcus mutans

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C:Accession: A42400

R:Russell, R.R.; Aduse-Opoku, J.; Sutcliffe, I.C.; Tao, L.; Ferretti, J.J.

J. Biol. Chem. 267, 4631-4637, 1992

A:Title: A binding protein-dependent transport system in Streptococcus mutans responsible

A:Reference number: A42400; MUID:92165821; PMID:1537846

A:Accession: A42400

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-278 <RUS>

A>Note: sequence extracted from NCBI backbone (NCBI:83887, NCBI:83888)

Query Match

Best Local Similarity 81.6%; Score 31; DB 2; Length 278;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTY 7

Db 72 KDITTY 78

RESULT 12

E69226

hypothetical protein MTH944 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: E69226

R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, R.; Dubois, J.; Aldredge, T.;

Oliver, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: E69226

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-279 <MTH>

A:Cross-references: GB:AE000869; GB:AE000666; NID:g2622042; PIDN:AAB85442.1; PID:g262204

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH944

Query Match

Best Local Similarity 81.6%; Score 31; DB 2; Length 279;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RDITTY 7

Db 54 RDITTY 60

RESULT 13

G69501

UDP-glucose 4-epimerase (gale-2) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C:Accession: G69501

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Cooney, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uitterlisch, T.; Cotton, M.D.; Springs, T.; Artlich, P.; Kalne, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: G69501

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-307 <KLE>

A:Cross-references: GB:AE000963; GB:AE000782; NID:g2689286; PIDN:AAB92234.1; PID:g264

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo

Query Match

Best Local Similarity 81.6%; Score 31; DB 2; Length 307;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RDITTY 7

Db 122 KDITTY 128

RESULT 14

S44133

histidine kinase - Lactococcus lactis

C:Species: Lactococcus lactis

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999

C:Accession: S44133

R:Engelke, G.; Glauski-Eckel, Z.; Klesau, P.; Siegers, K.; Hammelmann, M.; Enflen, K.

submitted to the EMBL Data Library, December 1993

A:Description: Regulation of nisin biosynthesis and immunity in L. lactis 6F3.

A:Reference number: S44130

A:Accession: S44133

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-447 <ENG>

A:Cross-references: EMBL:K76884; NID:g473019; PIDN:CA54212.1; PID:g473023

Query Match

Best Local Similarity 81.6%; Score 31; DB 2; Length 447;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DITTY 7

Db 332 DITTY 337

RESULT 15

E87397

transglycosylase, probable CCL194 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-Mar-2003

C:Accession: E87397

R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko

pro, Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87397

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-695 <STO>

A:Cross-references: GB:AE005673; NID:g13422517; PIDN:AAK23177.1; GSPDB:GN00148

C:Genetics:

A:Gene: CCL194

C:Superfamily: soluble lytic transglycosylase

Query Match

Best Local Similarity 81.6%; Score 31; DB 2; Length 695;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RDITTY 7

Db 401 RHITTY 407

Search completed: October 7, 2003, 19:21:52

Job time : 4.6069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:03:46 ; Search time 1.42414 Seconds

(without alignments)
231.148 Million cell updates/sec

Title: US-09-988-013a-4_COPY_99_105

Perfect score: 38

Sequence: 1 RDITTFY 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

127863

Total number of hits satisfying chosen parameters:

127863

Minimum DB seq length: 0

127863

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

127863

Maximum Match 100%

127863

Database : SwissProt_41.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	86.8	406	1	ARG1_HUMAN
2	33	86.8	414	1	ARG1_MOUSE
3	33	86.8	415	1	ARG1_RAT
4	32	84.2	377	1	NCAP_FITV
5	32	84.2	398	1	IF41_RABIT
6	32	84.2	406	1	IF41_HUMAN
7	32	84.2	407	1	IF42_MOUSE
8	32	84.2	407	1	IF42_HUMAN
9	31	81.6	278	1	MSMR_STRMO
10	31	81.6	447	1	NISK_LACIA
11	31	81.6	6359	1	BACC_BACLI
12	30	78.9	166	1	PAL_PSEPK
13	30	78.9	420	1	PUR2_STRSU
14	29	76.3	203	1	RASL_RHTRA
15	29	76.3	302	1	PCPB_PSEPA
16	29	76.3	357	1	ACT1_OXYFA
17	29	76.3	357	1	ACT1_OXYFA
18	29	76.3	375	1	ACT1_OXYNO
19	29	76.3	375	1	ACT1_OXYNO
20	29	76.3	375	1	ACT1_OXYNO
21	29	76.3	375	1	ACT1_OXYNO
22	29	76.3	428	1	GSA_STAMU
23	29	76.3	428	1	GSA_STAMU
24	29	76.3	473	1	HYA2_MOUSE
25	29	76.3	473	1	HYA2_MOUSE
26	29	76.3	510	1	HYA2_RAT
27	29	76.3	588	1	CYTD_ECOTI
28	29	76.3	716	1	PBPB_BACSU
29	29	76.3	880	1	YE21_ARCFU
30	29	76.3	951	1	SEF8_HUMAN
31	29	76.3	971	1	SEC5_YEAST
32	29	76.3	1061	1	DPOL_ADE12
33	29	76.3	1886	1	POL_COTMY

34	28	73.7	102	1	YD63_MYCPN
35	28	73.7	260	1	MCGR_METVA
36	28	73.7	265	1	YB46_MYCPN
37	28	73.7	308	1	Y178_TREPA
38	28	73.7	477	1	MRE2_ECOTI
39	28	73.7	499	1	CGAM_DROME
40	28	73.7	570	1	HEMA_NDVA
41	28	73.7	571	1	HEMA_NDVA
42	28	73.7	571	1	HEMA_NDVA3
43	28	73.7	571	1	HEMA_NDVA
44	28	73.7	571	1	HEMA_NDVA
45	28	73.7	571	1	HEMA_NDVA

ALIGNMENTS

RESULT 1	ID	ARG1_HUMAN	STANDARD:	PRT:	406 AA.
AC	08N6T3	Q96KC4	Q96R02	Q9NSU3	Q9NWF6
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	ADP-ribosylation factor GTPase activating protein 1 (ADP-ribosylation factor 1 GTPase activating protein) (ARF1 GAP) (ARF1-directed GTPase-activating protein) (GAP protein).				
OS	ARF1 GAP OR ARF1GAP.				
OC	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
NCBI_TaxID	9606;				
[1]	SEQUENCE FROM N.A. (ISOFORM 1).				
RA	Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,				
RA	Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,				
RA	Nakamura Y., Nagahara K., Masuho Y., Sasaki N.,				
RT	"NEBO human cDNA sequencing project."				
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
[2]	SEQUENCE FROM N.A.				
RA	MEDLINE-21636749; PubMed-11780052;				
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,				
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagdely C.L.,				
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,				
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,				
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,				
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,				
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,				
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,				
RA	Ellington A.G., Franklin J.A., Fraser A., French L., Garner P.,				
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,				
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,				
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,				
RA	Key M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,				
RA	Lehvasaisto M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,				
RA	Marsh V.L., Martin S.L., McConachie I.J., McKay K., McMurtry A.A.,				
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,				
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,				
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,				
RA	Rice C.M., Ross M.T., Scott C.E., Sehara H.K., Shownkeen R., Sims S.,				
RA	Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,				
RA	Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,				
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,				
RA	Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,				
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,				
RT	Rogers J.,				
RL	"The DNA sequence and comparative analysis of human chromosome 20.";				
[3]	Nature 414:865-871(2001).				
RA	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE-Fetal brain;				

RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner G., Shennan C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusla K., Farmer A.A., Rubin G.M., Hong L.,
RA Straptlen M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallat D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smolius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 240-406 FROM N.A. (ISOFORM 1).
RP TISSUE=Testis;
RC Duesterhoeft A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE OF 299-406 FROM N.A.
RP TISSUE=Fetal brain;
RT Ueki N.;
RT "HRI NTF human fetal brain cDNA project.";
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GTPase-activating protein (GAP) for the ADP ribosylation
CC factor 1 (ARF1). Involved in membrane trafficking and /or vesicle
CC transport. Promotes hydrolysis of the ARF1-bound GTP and thus, is
CC required for the dissociation of coat proteins from Golgi-derived
CC membranes and vesicles. A prerequisite for vesicle's fusion with
CC target compartment. Probably regulates ARF1-mediated transport via
CC its interaction with the KDELR proteins and RNP24. Overexpression
CC induces the redistribution of the entire Golgi complex to the
CC endoplasmic reticulum, as when ARF1 is deactivated. Its activity
CC is stimulated by phosphoinositides and inhibited by
CC phosphatidylinositol (By similarity).
CC -1- SUBUNIT: Interacts with ARF1. Interacts with the COPI coat
CC proteins, KDELR1 and RNP24. The interaction with RNP24 inhibits
CC the GAP activity (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; associates with the Golgi
CC complex (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms-2:
CC Name=1;
CC IsoId=Q8N6T3-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8N6T3-2; Sequence=VSP_000298; VSP_000299;
CC Note=no experimental confirmation available;
CC -1- DOMAIN: The region downstream of ARF-GAP domain is essential to
CC GAP activity in vivo. This region may be required for its
CC targeting to Golgi membranes (By similarity).
CC -1- SIMILARITY: Contains 1 Arf-GAP domain.
CC -1- CAUTION: Ref.4 sequence differs from that shown in that it seems
CC to include intronic sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AK021629; BAA91796.1; ALT. INT.
CC EMBL: AK027268; BAB55009.1; ALT. INT.
CC EMBL: AK027441; BAB55113.1; ALT. INT.

DR EMBL: AL121827; CAC36116.1; -;
DR EMBL: BC028233; AAH28233.1; -;
DR EMBL: AL137744; CAB70901.1; ALT-SEQ.
DR EMBL: AB015340; BAB8117.1; -;
DR Genew: HGNC:15852; ARFGAP1.
DR InterPro: IPR001164; hRIP_1like.
DR Pfam: PF01412; ArfGAP.1.
DR PRINTS: PR00405; REVINTACTING.
DR SMART: SM00105; ArfGAP.1.
DR PROSITE: PS50115; ARFGAP.1.
KM Transport: Protein transport; Golgi stack;
KM Zinc-finger: Alternative splicing; Polymorphism.
FT DOMAIN 7 124 ARF-GAP.
FT ZN-FING 22 45 C4-TYPE.
FT VARSPIC 239 239 K -> KFGCHKQOPEP (In isoform 2).
FT VARSPIC 279 280 /FTID-VSP_000298.
FT VARSPIC 279 280 Missing (In isoform 2).
FT VARIANT 184 184 /FTID-VSP_000299.
FT VARIANT 184 184 V -> M (In dbSNP:2273499).
FT CONFLICT 274 274 /FTID-VAR_015187.
FT CONFLICT 274 274 O -> R (In REF. 1; BAB55009).
SQ SEQUENCE 406 AA; 44668 MW; CAA1828DE660621 CRC64;

Qy Query Match 86.8%; Score 33; DB 1; Length 406;
Db Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RDITTFY 7
Db 288 RDYTFYF 294

RESULT 2
ARGL_MOUSE STANDARD; PRT; 414 AA.
ID 09EPY9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ADP-ribosylation factor GTPase activating protein 1 (ADP-ribosylation
DE factor 1 GTPase activating protein) (ARF1 GAP) (ARF1-directed GTPase-
DE activating protein) (GAP protein).
GN ARFGAP1 OR ARFGAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Venkateswaru K.;
RT "Cloning and functional characterisation of mouse ARFGAP.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GTPase-activating protein (GAP) for the ADP ribosylation
CC factor 1 (ARF1). Involved in membrane trafficking and /or vesicle
CC transport. Promotes hydrolysis of the ARF1-bound GTP and thus, is
CC required for the dissociation of coat proteins from Golgi-derived
CC membranes and vesicles. A prerequisite for vesicle's fusion with
CC target compartment. Probably regulates ARF1-mediated transport via
CC its interaction with the KDELR proteins and RNP24. Overexpression
CC induces the redistribution of the entire Golgi complex to the
CC endoplasmic reticulum, as when ARF1 is deactivated. Its activity
CC is stimulated by phosphoinositides and inhibited by
CC phosphatidylinositol (By similarity).
CC -1- SUBUNIT: Interacts with ARF1. Interacts with the COPI coat
CC proteins, KDELR1 and RNP24. It is probably a component of the COPI
CC coat protein complex. The interaction with RNP24 inhibits the GAP
CC activity (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; associates with the Golgi
CC complex (By similarity).
CC -1- DOMAIN: The region downstream of ARF-GAP domain is essential to
CC GAP activity in vivo. This region may be required for its
CC targeting to Golgi membranes (By similarity).
CC -1- SIMILARITY: Contains 1 Arf-GAP domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ401461; CAC18721.1; -
DR MGD: MGI:218359; Arfgap1.
DR InterPro: IPR001164; hRIP-like.
DR Pfam: PF01412; ArfGap; 1.
DR PRINTS: PR00405; REVINTRACTING.
DR SMART: SM00105; ArfGap; 1.
DR PROSITE: PS50115; ARFGAP; 1.
KW Transport: Protein transport; GTPase activation; Golgi stack;
KW Zinc-finger.
FT DOMAIN 7 124 ARF-GAP.
FT ZN_FING 22 45 C4-TYPE.
SQ SEQUENCE 414 AA; 45269 MW; 0A050D29DE90BC62 CRC64;
Query Match 86.88; Score 33; DB 1; Length 414;
Best Local Similarity 71.48; Pred. No. 8.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RDIITFY 7
Db 288 RDVTTF 294

RESULT 3
ARG1_RAF
ID ARG1_RAF STANDARD; PRT; 415 AA.
AC Q62848;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADP-ribosylation factor GTPase activating protein 1 (ADP-ribosylation
DE factor 1 GTPase activating protein) (ARF1 GAP) (ARF1-directed GTPase-
DE activating protein) (GAP protein).
GN ARFGAP1 OR ARF1GAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, INTERACTION WITH ARF1, AND SEQUENCE OF
RP 82-95: 140-160; 183-191; 246-253; 264-278; 297-306 AND 361-375.
RC TISSUE=Liver;
RX MEDLINE=95197515; PubMed=7890632;
RA Maier V., Cukierman E., Rotman M., Admon A., Cassel D.;
RT "ADP-ribosylation factor-directed GTPase-activating protein.
RT Purification and partial characterization.";
RL J. Biol. Chem. 270:5232-5237(1995).
[2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), CHARACTERIZATION, AND
RP MUTAGENESIS OF CYS-22 AND CYS-25.
RC TISSUE=Liver;
RX MEDLINE=96123274; PubMed=8533093;
RA Cukierman E., Huber I., Rotman M., Cassel D.;
RT "The ARF1 GTPase-activating protein: zinc finger motif and Golgi
RT complex localization.";
RL Science 270:1999-2002(1995).
[3]
RN INTERACTION WITH KDELR1.
RP MEDLINE=98070308; PubMed=9405360;
RA Aoe T., Cukierman E., Lee A., Cassel D., Peters P.J., Hsu V.W.;
RT "The KDELR receptor, ERD2, regulates intracellular traffic by
RT recruiting a GTPase-activating protein for ARF1.";
RL EMBO J. 16:7305-7316(1997).
[4]
RP CHARACTERIZATION, AND MUTAGENESIS OF CYS-22.

RX MEDLINE=98406131; PubMed=9733781;
RA Huber I., Cukierman E., Rotman M., Aoe T., Hsu V.W., Cassel D.;
RT "Requirement for both the amino-terminal catalytic domain and a
RT noncatalytic domain for in vivo activity of ADP-ribosylation factor
RT GTPase-activating protein.";
RL J. Biol. Chem. 273:24786-24791(1998).
[5]
RN INTERACTION WITH RNP24.
RX MEDLINE=21629435; PubMed=11748249;
RA Lanoix J., Ouwendijk J., Stark A., Szafer E., Cassel D., Dejgaard K.,
RA Weiss M., Nilsson T.;
RT "Sorting of Golgi resident proteins into different subpopulations of
RT COP1 vesicles: a role for ArfGap1.";
RL J. Cell Biol. 155:1199-1212(2001).
[6]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF CATALYTIC DOMAIN IN COMPLEX
RP WITH HUMAN ARF1.
RX MEDLINE=99200399; PubMed=10102276;
RA Goldberg J.;
RT "Structural and functional analysis of the ARF1-ARFGAP complex reveals
RT a role for coatomer in GTP hydrolysis.";
RL Cell 96:893-902(1999).
[7]
CC -1- FUNCTION: GTPase-activating protein (GAP) for the ADP ribosylation
CC factor 1 (ARF1). Involved in membrane trafficking and/or vesicle
CC transport. Promotes hydrolysis of the ARF1-bound GTP and thus, is
CC required for the dissociation of coat proteins from Golgi-derived
CC membranes and vesicles, a prerequisite for vesicle's fusion with
CC target compartment. Probably regulates ARF1-mediated transport via
CC its interaction with the KDELR proteins and RNP24. Overexpression
CC induces the redistribution of the entire Golgi complex to the
CC endoplasmic reticulum, as when ARF1 is deactivated. Its activity
CC is stimulated by phospholipids and inhibited by
CC phosphatidylcholine.
CC -1- SUBUNIT: Interacts with ARF1. Interacts with the COP1 coat
CC proteins, KDELR1 and RNP24. It is probably a component of the COP1
CC coat protein complex. The interaction with RNP24 inhibits the GAP
CC activity.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; associates with the Golgi
CC complex.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Comment-Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q62848-1; Sequence=Displayed;
CC Name=2; Synonyms=K15;
CC IsoId=Q62848-2; Sequence=VSP_000300;
CC Name=3; Synonyms=Z5;
CC IsoId=Q62848-3; Sequence=VSP_000301, VSP_000302;
CC -1- TISSUE SPECIFICITY: Widely expressed. Highly expressed in brain and
CC liver.
CC -1- DOMAIN: The region downstream of ARF-GAP domain is essential to
CC GAP activity in vivo. This region may be required for its
CC targeting to Golgi membranes.
CC -1- SIMILARITY: Contains 1 Arf-GAP domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U35776; AAC52337.1; -
DR InterPro: IPR001164; hRIP-like.
DR Pfam: PF01412; ArfGap; 1.
DR PRINTS: PR00405; REVINTRACTING.
DR SMART: SM00105; ArfGap; 1.
DR PROSITE: PS50115; ARFGAP; 1.
KW Transport: Protein transport; GTPase activation; Golgi stack;
KW Zinc-finger; Alternative splicing.
FT DOMAIN 7 124 ARF-GAP.

FT 2N.FING 22 45 C4-TYPE.
 FT VARSPPLIC 21 57 Missing (in isoform 2).
 FT VARSPPLIC 280 283 /FTid=VSP_000300.
 FT VARSPPLIC 284 415 /FTid=VSP_000301.
 FT VARSPPLIC 284 415 Missing (in isoform 3).
 FT MTTAGEN 22 22 /FTid=VSP_000302.
 FT MTTAGEN 25 25 C->A: LOSS OF GAP ACTIVITY.
 FT MTTAGEN 25 25 C->A: LOSS OF GAP ACTIVITY.
 SQ SEQUENCE 415 AA; 45442 MW; 90015417E3E717BB CRC64;

Query Match 86.8%; Score 33; DB 1; Length 415;
 Best Local Similarity 71.4%; Pred. No. 8.7;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
 Db 288 RDVTFY 294

RESULT 4
 NCAP_FIPV STANDARD; PRT; 377 AA.
 ID NCAP_FIPV P25909;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DN Nucleocapsid protein (N structural protein).
 GN N.

OS Feline infectious peritonitis virus (strain 79-1146) (FIPV).
 OC Viruses; SARS positive-strand viruses, no DNA stage. Nidovirales;
 OC Coronaviridae; Coronaviruses.
 NC NCBL_TaxID=33734;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91134997; PubMed=1847259;
 RA Vennema H., de Groot R.J., Harbour D.A., Horzinek M.C.,
 RA Spaan W.J.M.;
 RT Primary structure of the membrane and nucleocapsid protein genes of
 RT feline infectious peritonitis virus and immunogenicity of recombinant
 RT vaccinia viruses in kittens.";
 RL Virology 181:327-335(1991).
 CC -1- SIMILARITY: Belongs to the coronavirus nucleocapsid protein
 CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

CC EMBL: X56496; CAA39851.1; -
 CC PIR: B38498; VHIH79.
 DR InterPro: IPR001218; Corona_nucleocap.
 DR Pfam: PF00937; Corona_nucleoca. 1.
 KW Nucleocapsid.
 SQ SEQUENCE 377 AA; 42745 MW; D76382AE688059B CRC64;

Query Match 84.2%; Score 32; DB 1; Length 377;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DITTFY 7
 Db 241 DVTFY 246

RESULT 5
 IF41_RABIT STANDARD; PRT; 398 AA.
 ID IF41_RABIT P29562;
 AC P29562;

DT 01-APR-1993 (Rel. 25, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Eukaryotic initiation factor 4A-1 (eIF-4A-1) (Fragment).
 GN EIF4A1 OR EIF4A.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NC NCBL_TaxID=9986;
 RN [1]

RP SEQUENCE OF 5-398 FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE=91053155; PubMed=2241157;
 RA Conroy S.C., Deyer T.E., Owens C.L., Merrick W.C.;
 RT Characterization of the 46,000-dalton subunit of eIF-4F.";
 RL Arch. Biochem. Biophys. 282:363-371(1990).
 CC -1- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
 CC PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS REQUIRED AS A
 CC SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING TO RIBOSOME. IT IS AN
 CC ATP-DEPENDENT SINGLE STRANDED DNA-BINDING PROTEIN WITH A SEQUENCE-
 CC INDEPENDENT UNWINDING ACTIVITY (HELICASE).

CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC HSP; Q58083; 1HV8.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD_box.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 KW Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding;
 KW DNA-binding; Helicase.
 FT NON_TER 1
 FT NP_BIND 68 75 ATP (BY SIMILARITY).
 FT SITE 174 177 DEAD BOX.
 SQ SEQUENCE 398 AA; 45291 MW; 8CE1C18DE7AC7D37 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 398;
 Best Local Similarity 85.7%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
 Db 377 RDITTFY 383

RESULT 6
 IF41_HUMAN STANDARD; PRT; 406 AA.
 ID IF41_HUMAN P04765; O61516;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Eukaryotic initiation factor 4A-1 (eIF-4A-1) (eIF4A-1).
 GN EIF4A1 OR EIF4A OR DDXX2A.
 OS Homo sapiens (Human), and
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBL_TaxID=9606, 10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=93261841; PubMed=8493113;
 RA Kim N.-S., Kato T., Abe N., Kato S.;
 RT "Nucleotide sequence of human cDNA encoding eukaryotic initiation
 RT factor 4A1.";
 RL Nucleic Acids Res. 21:2012-2012(1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RA SPECIES=Human; TISSUE=Bone marrow;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman A., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE;
 RX MEDLINE=86041884; PubMed=3840589;
 RA Nielsen P.J., McMaster G.K., Trachsel H.,
 RT "Cloning of eukaryotic protein synthesis initiation factor genes:
 RT isolation and characterization of cDNA clones encoding factor
 RT eIF-4A.";
 RL Nucleic Acids Res. 13:6867-6880(1985).
 [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE;
 RX MEDLINE=89108007; PubMed=32151517;
 RA Reddy N.S., Roth W.W., Bragg P.W., Mahba A.J.,
 RT "Isolation and mapping of a gene for protein synthesis initiation
 RT factor 4A and its expression during differentiation of murine
 RT erythroleukemia cells.";
 RL Gene 70:231-243(1988).
 [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE; STRAIN=129;
 RA Miyashita A., Shimizu N., Nakajima T., Odani S., Kuwano R.,
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 [6]
 RP MUTAGENESIS.
 RC SPECIES-MOUSE;
 RX MEDLINE=92331617; PubMed=1378397;
 RA Pause A., Sonnenberg N.,
 RT "Mutational analysis of a DEAD box RNA helicase: the mammalian
 RT translation initiation factor eIF-4A.";
 RL EMBO J. 11:2643-2654(1992).
 CC -1- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
 CC PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS REQUIRED AS A
 CC SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING TO RIBOSOME. IT IS AN
 CC ATP-DEPENDENT SINGLE STRANDED DNA-BINDING PROTEIN WITH A SEQUENCE-
 CC INDEPENDENT UNWINDING ACTIVITY (HELICASE).
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D13748; BAA02897.1; -;
 DR EMBL: BC009585; AAH09585.1; -;
 DR EMBL: X03039; CAA26842.1; ALT_INIT.
 DR EMBL: X03039; CAA26843.1; ALT_INIT.
 DR EMBL: X03040; CAA26845.1; ALT_INIT.
 DR EMBL: X03040; CAA26846.1; ALT_INIT.
 DR EMBL: L36611; AAA50407.1; -;
 DR EMBL: M22873; AAA50407.1; JOINED.

DR EMBL: L36608; AAA50407.1; JOINED.
 DR EMBL: L36609; AAA50407.1; JOINED.
 DR EMBL: L36610; AAA50407.1; JOINED.
 DR EMBL: AB011595; BAA25075.1; -;
 DR PIR: J50039; FIMS4A.
 DR PIR: S33681; S33681.
 DR HSSP: Q58083; IHV8.
 DR Genew: HGNC:3282; EIF4A1.
 DR GK: P04765; -;
 DR MIM: 602641; -;
 DR MGD: MGI:95303; EIf4a1.
 DR GO: GO:0008306; C:eukaryotic translation initiation factor 4 . . .; TAS.
 DR GO: GO:0003729; F:mRNA binding activity; TAS.
 DR GO: GO:0006441; P:binding to mRNA cap; TAS.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD box.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICC; 1.
 DR PROSITE: PS00039; DEAD.ATP.HELICASE; 1.
 DR Initiation factor: Protein biosynthesis; ATP-binding; RNA-binding;
 KW DNA-binding; Helicase.
 FT NP_BIND 76 83 ATP (BY SIMILARITY).
 FT SITE 182 185 DEAD BOX.
 FT CONFLICT 300 300 S -> C (IN REF. 4).
 SQ SEQUENCE 406 AA; 46154 MW; 6EF89939F3045420 CRC64;
 QY 1 RDIITFY 7 84.2% Score 32; DB 1; Length 406;
 Db 385 RDIITFY 391 Best Local Similarity 85.7% Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 RESULT 7
 IF42_HUMAN STANDARD; PRT; 407 AA.
 AC Q14240; O96B90; -;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Eukaryotic initiation factor 4A-II (eIF-4A-II).
 GN EIF4A2 OR DDXXB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=96103476; PubMed=8521730;
 RA Sudo K., Takahashi E., Nakamura Y.,
 RT "Isolation and mapping of the human EIF4A2 gene homologous to the
 RT murine protein synthesis initiation factor 4A-II gene EIf4a2.";
 RL Cytogenet. Cell Genet. 71:385-388(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smillius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
 CC PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS REQUIRED AS A
 CC SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING TO RIBOSOME. IT IS AN
 CC ATP-DEPENDENT SINGLE STRANDED DNA-BINDING PROTEIN WITH A SEQUENCE-
 CC INDEPENDENT UNWINDING ACTIVITY (HELICASE).
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chromosome/genet/EIF4A2ID262.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D30655; BAA06336.1; -;
 CC DR EMBL: BC015842; AAH15842.1; -;
 CC DR HSSP: Q58083; I1V8.
 CC DR Genew: HGNC:3284; EIF4A2.
 CC DR GK: 014240; -;
 CC DR KIM: 601102; -;
 CC DR GO: GO:0008304; C:eukaryotic translation initiation factor 4 . . .; TAS.
 CC DR GO: GO:0003743; F:translation initiation factor activity; TAS.
 CC DR GO: GO:0006446; P:regulation of translational initiation; TAS.
 CC DR InterPro: IPR001410; DEAD.
 CC DR InterPro: IPR000629; DEAD_box.
 CC DR InterPro: IPR001650; Helicase_C.
 CC DR Pfam: PF00270; DEAD; 1.
 CC DR Pfam: PF00271; Helicase_C; 1.
 CC DR SMART: SM00487; DEXDC; 1.
 CC DR SMART: SM00490; HELIC; 1.
 CC DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 CC DR Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding;
 CC KW DNA-binding; Helicase.
 CC KM NP_BIND: 77
 CC FT SITE: 183 186 ATP (BY SIMILARITY).
 CC FT SITE: 27 27 DEAD BOX.
 CC FT CONFLICT: 212 213 N -> S (IN REF. 2).
 CC FT CONFLICT: 212 213 FA -> LT (IN REF. 2).
 CC SO SEQUENCE 407 AA; 46394 MW; 046BA69D25ED5A87 CRC64;
 CC
 CC Query Match 84.2%; Score 32; DB 1; Length 407;
 CC Best Local Similarity 85.7%; Pred. No. 14;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 RDITFEY 7
 CC Db 386 RDITFEY 392
 CC
 CC RESULT 8
 CC IF42_MOUSE STANDARD; PRT; 407 AA.
 CC ID IF42_MOUSE Q61513; Q61514;
 CC AC P10630; Q61513; Q61514;
 CC DT 01-JUL-1989 (Rel. 11, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Eukaryotic initiation factor 4A-II (EIF-4A-II) (EIF4A-II).
 CC GN EIF4A2 OR DDX2B.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI-TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT SER-9 INS.
 RX MEDLINE-8832898; PubMed-3046931;
 RA Nielsen P.J., Trachsel H.;
 RT "The mouse protein synthesis initiation factor 4A gene family
 RT includes two related functional genes which are differentially
 RT expressed".
 RL EMBO J. 7:2097-2105(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-Liver;
 RA Nielsen P.J.;
 RN Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE OF 291-323 FROM N.A.
 RC TISSUE-Liver;
 RX MEDLINE-97248616; PubMed-9092667;
 RA Selvaratnam N., Joost O.H., Haas E.S., Brown J.W., Galvin N.J.,
 RA Ellicott G.L.;
 RT "Intracellular localization and unique conserved sequences of three
 RT small nucleolar RNAs".
 RL Nucleic Acids Res. 25:1591-1596(1997).
 CC -1- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
 CC PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS REQUIRED AS A
 CC SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING TO RIBOSOME. IT IS AN
 CC ATP-DEPENDENT SINGLE STRANDED DNA-BINDING PROTEIN WITH A SEQUENCE-
 CC INDEPENDENT UNWINDING ACTIVITY (HELICASE).
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X12507; CAA31025.1; -;
 CC DR EMBL: X14422; CAA32585.1; -;
 CC DR EMBL: X56953; CAA40269.1; -;
 CC DR EMBL: X56953; CAA40268.1; -;
 CC DR EMBL: U64706; AAC53180.1; -;
 CC DR PIR: S00985; S00985.
 CC DR HSSP: Q58083; I1V8.
 CC DR MGD: MGI:106906; Elf4a2.
 CC DR InterPro: IPR001410; DEAD.
 CC DR InterPro: IPR000629; DEAD_box.
 CC DR InterPro: IPR001650; Helicase_C.
 CC DR Pfam: PF00270; DEAD; 1.
 CC DR Pfam: PF00271; Helicase_C; 1.
 CC DR SMART: SM00487; DEXDC; 1.
 CC DR SMART: SM00490; HELIC; 1.
 CC DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 CC DR Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding;
 CC KW DNA-binding; Helicase.
 CC KM NP_BIND: 77
 CC FT SITE: 183 186 ATP (BY SIMILARITY).
 CC FT SITE: 9 9 DEAD BOX.
 CC FT VARIANT: 200 201 IF -> RV (IN REF. 1).
 CC FT CONFLICT: 294 297 HARD -> QATY (IN REF. 1).
 CC FT CONFLICT: 387 387 D -> H (IN REF. 1).
 CC FT CONFLICT: 403 403 V -> G (IN REF. 1).
 CC SO SEQUENCE 407 AA; 46402 MW; FAAVD3BA9D8C6DA0 CRC64;
 CC
 CC Query Match 84.2%; Score 32; DB 1; Length 407;
 CC Best Local Similarity 85.7%; Pred. No. 14;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 RDITFEY 7
 CC Db 386 RDITFEY 392

```

RESULT 9
MSMR_STRMU STANDARD: PRT: 278 AA.
ID MSMR_STRMU
AC Q00753;
DT 01-JUN-1994 (Rel. 29, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mem operon regulatory protein.
GN MSMR OR SMU.876.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1309;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ingbritt;
RX MEDLINE=92165821; PubMed=1537846;
RA Russell R.R.B., Aduse-Opoku J., Sutcliffe I.C., Tao L.,
RA Ferretti J.J.;
RT "A binding protein-dependent transport system in Streptococcus mutans
RT responsible for multiple sugar metabolism."
RL J. Biol. Chem. 267:4631-4637(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=0A159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans 0A159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: REGULATORY PROTEIN FOR THE MSM OPERON FOR MULTIPLE
CC SUGAR METABOLISM. ACTIVATES THE TRANSCRIPTION OF THE MSMERG,
CC AGS, DEXA AND GFTA GENES.
CC -1- SIMILARITY: BELONGS TO THE ARAC/XLIS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M77351; AAA26932.1; -.
DR EMBL: AB014929; AAN58591.1; -.
DR InterPro: IPR000005; HTHARAC.
DR Pfam: PF00165; HTH_ARAC; 2.
DR PRINTS: PR00032; HTHARAC.
DR SMART: SM00342; HTH_ARAC; 1.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
KM Transcription regulation; DNA-binding; Activator; Complete proteome.
FT DNAS_BIND 192 211 L->V (IN REF. 1).
FT CONFLICT 4 4 SSKAENLKSSSLFKR -> RVQKPKFLSSQAVLNL
FT CONFLICT 94 110 (IN REF. 1).
SQ SEQUENCE 278 AA; 31967 MW; 9A8E28ABF6CCCEA CRC64;
Query Match 81.6%; Score 31; DB 1; Length 278;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 RDITFFY 7
DB 72 KDISTFY 78

```

```

ID NISK_LACIA STANDARD: PRT: 447 AA.
AC P42707;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nisin biosynthesis sensor protein nisk (EC 2.7.3.-).
GN NISK.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OC NCBI_TaxID=1360;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=63;
RX MEDLINE=94213458; PubMed=8161176;
RA Engleke G., Gutowski-Eckel Z., Klesau P., Siegers K.,
RA Hammelmann M., Ertlan R.-D.;
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
RT 63."
RL Appl. Environ. Microbiol. 60:814-825(1994).
CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM NISK/NISR
CC INVOLVED IN THE REGULATION OF THE BIOSYNTHESIS OF LANTIBIOTIC
CC NISIN. NISK MAY FUNCTION AS A MEMBRANE-ASSOCIATED PROTEIN
CC KINASE THAT PHOSPHORYLATES NISR IN RESPONSE TO ENVIRONMENTAL
CC SIGNALS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X76884; CA54212.1; -.
DR PIR: S44133; S44133.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR005467; His_Kinase.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; HisKA_1.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00388; HisKA_1.
DR PROSITE: PS00109; HIS_KIN; 1.
KM Sensory transduction; Transferase; Kinase; Transmembrane;
KV Phosphorylation.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT DOMAIN 235 447 HISTIDINE KINASE.
FT MOD_RES 238 238 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 305 305 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 447 AA; 51319 MW; 47E10533F5A9685D CRC64;
Query Match 81.6%; Score 31; DB 1; Length 447;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 DITFFY 7
DB 332 DITFFY 337

```

```

RESULT 11
BACC_BACLI STANDARD: PRT: 6359 AA.
ID BACC_BACLI
AC O68008;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bacitracin synthetase 3 (Ba3) [Includes: Atp-dependent isoleucine
DE adenylation (IleA) (isoleucine activase); Atp-dependent D-phenylalanine
DE adenylation (D-PheA) (D-phenylalanine activase); Atp-dependent histidine

```

DE adenylation (HSA) (Histidine activase); ATP-dependent D-aspartate
 DE adenylation (D-AspA) (D-aspartate activase); ATP-dependent D-aspartate
 DE adenylation (AsnA) (Asparagine activase); Aspartate racemase
 DE (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]
 DE (EC 5.1.1.11).
 GN BACC.
 OS Bacillus licheniformis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1402.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10716;
 RX MEDLINE=98089193; PubMed=9427658;
 RA Konz D., Klens A., Schoegendorfer K., Marahel M.A.;
 RT "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC
 RT 10716: molecular characterization of three multi-modular peptide
 RT synthetases.";
 RT Chem. Biol. 4:927-937(1997).
 CC -1- FUNCTION: INDUCES PEPTIDE SYNTHESIS, ACTIVATES AND INCORPORATES
 CC FIVE AMINO ACIDS, FORMS A THIAZOLINE RING BETWEEN THE FIRST TWO
 CC AMINO ACIDS AND INCORPORATES A D-GLUTAMINE IN THE FOURTH POSITION.
 CC -1- CATALYTIC ACTIVITY: L-aspartate -> D-aspartate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine -> AMP + diphosphate + D-
 CC phenylalanine.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (PORENTIAL).
 CC -1- PATHWAY: Cyclic peptide antibiotic bacitracin biosynthesis.
 CC -1- SUBUNIT: LARGE MULTISUBUNIT COMPLEX OF BA1, BA2 AND BA3.
 CC -1- DOMAIN: CONSISTS OF FIVE MODULES WITH TWO EPIMERIZATION DOMAINS IN
 CC THE SECOND AND FOURTH MODULES, AND A PUTATIVE C-TERMINAL
 CC THIOESTERASE DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO
 CC THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS
 CC RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOALATION, CONDENSATION
 CC (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND
 CC N-METHYLATION (OPTIONAL).
 CC -1- MISCELLANEOUS: BACITRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC
 CC DODECAPETIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST
 CC ABUNDANT IS BACITRACIN A, A BRANCHED CYCLIC DODECAPETIDE. IT
 CC CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (ILE-CYS-LEU-D-
 CC GLU-ILE) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION
 CC PRODUCT AND A C-TERMINAL HEPTAPEPTIDE RING (LYS-D-ORN-ILE-D-PHE-
 CC HIS-D-ASP-ASN), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE C-
 CC TERMINAL ASN IS BOUND TO THE EPSILON-AMINO GROUP OF LYSINE. IT
 CC CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,
 CC PHE-9, AND ASP-11).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: Contains 5 acyl carrier domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF007865; AAC06348.1; .
 CC DR PIR: T31679; T31679.
 CC DR HSSP: P14687; 1AMU.
 CC DR InterPro: IPR000873; AMP-bind.
 CC DR InterPro: IPR001242; Condensatn.
 CC DR InterPro: IPR006163; PP_bind.
 CC DR InterPro: IPR006162; Phantse-attach.
 CC DR InterPro: IPR001031; Thioesterase.
 CC DR Pfam: PF00501; AMP-binding; 5.
 CC DR Pfam: PF00668; Condensation; 7.
 CC DR Pfam: PF00550; PP-binding; 5.
 CC DR Pfam: PF00975; Thioesterase; 1.
 CC DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 4.
 CC DR PROSITE: PS00455; AMP-BINDING; 5.
 CC DR PROSITE: PS50075; ACP DOMAIN; 5.
 CC Lysase; Isomerase; Hydrolase; Antibiotic biosynthesis;

KW Phosphopantetheine: Multifunctional enzyme: Repeat.
 FT REPEAT 461 1034
 FT REPEAT 1517 2064
 FT REPEAT 2999 3570
 FT REPEAT 4047 4612
 FT REPEAT 5549 6129
 FT DOMAIN 966 1033
 FT DOMAIN 1998 2064
 FT DOMAIN 3502 3569
 FT DOMAIN 4544 4610
 FT DOMAIN 6052 6119
 FT BINDING 996 996
 FT BINDING 2028 2028
 FT BINDING 3532 3532
 FT BINDING 4574 4574
 FT BINDING 6082 6082
 SQ SEQUENCE 6359 AA; 722923 MW; 82A273C546253074 CXC64;
 Query Match 81.6%; Score 31; DB 1; Length 6359;
 Best Local Similarity 71.4%; Pred. No. 4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 IDITPEY 7
 Db 4305 RDVETPY 4311
 ID PAL_PSEPK STANDARD; PRT; 166 AA.
 ID PAL_PSEPK
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Peptidoglycan-associated lipoprotein precursor.
 GN PAL OR PalI OR OprL OR Pp123.
 OS Pseudomonas putida (strain KT2440), and
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488, 303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KT2440.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinl C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Kozaczek A.,
 RA Ueberbacher T., Rizzo M., Lee K., Kosack D., Mestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Helm S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RT Environ. Microbiol. 4:799-808(2002).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=mt-2.
 RC MEDLINE=96198174; PubMed=8626299;
 RX Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.L.;
 RT "The pseudomonas putida peptidoglycan-associated outer membrane
 RT lipoprotein is involved in maintenance of the integrity of the cell
 RT cell envelope.";
 RT J. Bacteriol. 178:1699-1706(1996).
 CC -1- FUNCTION: VERY STRONGLY ASSOCIATED WITH THE PEPTIDOGLYCAN (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor.
 CC -1- SIMILARITY: BELONGS TO THE OMPA FAMILY. PAL SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; A6016778; AAN66847.1; -
 DR EMBL; X74218; CAA52294.1; -
 DR PIR; S52308; S52308.
 DR TIGR; PPI223; -
 DR InterPro: IPR006664; Bac.OmpA.
 DR InterPro: IPR006665; OmpA/MotB.
 DR InterPro: IPR006690; OmpA-Like.
 DR Pfam; PF00691; OmpA; 1.
 DR PRINTS; PR01021; OMPADOMAIN.
 DR PRODOM; PD000930; OMPA/MotB; 1.
 DR PROSITE; PS00013; PROKAR.LIPOPROTEIN; 1.
 DR PROSITE; PS01068; OMPA; 1.
 KW Outer membrane; Signal; Lipoprotein; Complete proteome.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 166 PEPTIDOLYCAN-ASSOCIATED LIPOPROTEIN.
 FT LIPID 22 22 N-ACYL DIGLYCERIDE (BY SIMILARITY).
 FT DOMAIN 98 142 OMPA-LIKE
 SO SEQUENCE 166 AA; 17833 MW; 1735318143555A01 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 166;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DITTFY 7
 I I I I I
 DB 60 RAITTFY 66

RESULT 13
 PUR2_STRSU STANDARD; PRT; 420 AA.
 ID Q9F1S9;
 AC Q9F1S9;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phosphoribosylamine-glycine ligase (EC 6.3.4.13) (GARS) (glycinamide
 DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase).
 GN PUSD.
 OS Streptococcus suis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1307;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DAT1 / Serotype 2;
 RX MEDLINE=20576151; PubMed=1113943;
 RA Sekizaki T., Otani Y., Osaki M., Takamatsu D., Shimoi Y.;
 RT "Evidence for horizontal transfer of the SUSDAT1 restriction-
 RT modification genes to the Streptococcus suis genome.";
 RL J. Bacteriol. 183:500-511(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + 5-phospho-D-riboseylamine + glycine - ADP
 CC + phosphate + N(1)-(5-phospho-D-riboseyl)glycinamide.
 CC -1- PATHWAY: De novo purine biosynthesis; second step.
 CC -1- SIMILARITY: Belongs to the GARS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB045609; BAB20832.1; -
 DR HSSP; P15640; IGSO.
 DR HAMAP; MF_00138; -; 1.

DR InterPro: IPR000115; Gars.
 DR Pfam; PF01071; GARS; 1.
 DR Pfam; PF02842; GARS_B; 1.
 DR Pfam; PF02843; GARS_C; 1.
 DR Pfam; PF02844; GARS_N; 1.
 DR TIGRfams; TIGR00877; purD; 1.
 DR PROSITE; PS00184; GARS; 1.
 KW Purine biosynthesis; Ligase.
 SO SEQUENCE 420 AA; 45333 MW; 14E9C2D29CDF6B9F CRC64;

Query Match 78.9%; Score 30; DB 1; Length 420;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DITTFY 7
 I I I I I
 DB 355 DITTFY 360

RESULT 14
 RAS1_RHRA STANDARD; PRT; 203 AA.
 ID RAS1_RHRA
 AC P22278;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-like protein 1.
 GN RAS1.
 OS Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 CC Mucor.
 OX NCBI_TaxID=4841;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 1216B;
 RX MEDLINE=91061774; PubMed=1701021;
 RA Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
 RT "Expression of a gene family in the dimorphic fungus Mucor racemosus
 RT which exhibits striking similarity to human ras genes.";
 RL Mol. Cell. Biol. 10:6654-6663(1990).
 CC -1- ENZYME REGULATION: ALTERNATE BETWEEN AN INACTIVE FORM BOUND TO GDP
 CC AND AN ACTIVE FORM BOUND TO GTP. ACTIVATED BY A GUANINE
 CC NUCLEOTIDE-EXCHANGE FACTOR (GEF) AND INACTIVATED BY A GTPASE-
 CC ACTIVATING PROTEIN (GAP).
 CC -1- SUBCELLULAR LOCATION: Plasma membrane.
 CC -1- DEVELOPMENTAL STAGE: IN ALL DEVELOPMENTAL STAGES ANALYZED. ITS
 CC SIGNAL WAS MORE INTENSE IN SPORULATING MYCELION.
 CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M55175; AAA83378.1; -
 DR PIR; A36365; A36365.
 DR HSSP; P01112; IPL.
 DR InterPro: IPR003577; GTPase_Ras.
 DR InterPro: IPR001806; Ras_trnsmfmg.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRANG.
 DR SMART; SM00173; RAS; 1.
 DR TIGRfams; TIGR00231; small_gtp; 1.
 KW GTP-binding; Prenylation; Lipoprotein.
 FT NP_BIND 17 24 GTP (BY SIMILARITY).
 FT NP_BIND 64 68 GTP (BY SIMILARITY).
 FT NP_BIND 123 126 GTP (BY SIMILARITY).
 FT DOMAIN 39 47 EFFECTOR REGION (PROBABLE).
 FT LIPID 200 200 FARNESYL (BY SIMILARITY).

SO SEQUENCE 203 AA: 23236 MW: 52098F53F396A54 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 203;
Best Local Similarity 83.3%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DIFFY 7
:|||||
DB 99 DIFFY 104

RESULT 15

PCYB_PSEPA
ID PCYB_PSEPA STANDARD; PRT; 302 AA.
AC P22636;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protocatechuate 4,5-dioxygenase beta chain (EC 1.13.11.8) (4,5-PCD).
GN LIGB.
OS Pseudomonas paucimobilitis (Sphingomonas paucimobilitis).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Sphingomonas.
OX NCBI_TaxID=13689;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.
RC STRAIN-SYK-6;
RX MEDLINE=90236935; PubMed=2185230;
RA Noda Y., Nishikawa S., Shiozuka K.-I., Kadokura H., Nakajima H.,
RA Yoda K., Katayama Y., Morohoshi N., Harauchi T., Yamasaki M.;
RT "Molecular cloning of the protocatechuate 4,5-dioxygenase genes of
RT Pseudomonas paucimobilitis.";
RL J. Bacteriol. 172:2704-2709(1990).
CC -|- FUNCTION: RESPONSIBLE FOR THE AROMATIC RING FISSION OF
CC PROTocatechuate.
CC -|- CATALYTIC ACTIVITY: Protocatechuate + O(2) -> 4-carboxy-2-
CC hydroxymuconate semialdehyde.
CC -|- COFACTOR: FERROUS ION.
CC -|- SUBUNIT: COMPOSED OF TWO SUBUNITS (ALPHA AND BETA) IN A 1:1 RATIO.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M34835; AAA17728.1; -
DR PIR: B35271; B35271.
DR PDB: 1BAU; 27-AUG-99.
DR PDB: 1B0U; 04-MAY-99.
DR InterPro: IPR004183; LIGB.
DR Pfam: PF02900; LIGB; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxygenase; Iron;
KW 3D-structure.
SQ SEQUENCE 302 AA: 33292 MW: 0552B3B0E59702E5 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 302;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITTFY 7
||:||||
DB 258 RDITTFY 264

Search completed: October 7, 2003, 19:15:04
Job time : 2.42414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 19:04:06 ; Search time 6.97586 Seconds
(without alignments)
258.946 Million cell updates/sec

Title: US-09-988-013a-4_COPY_99_105
Perfect score: 38
Sequence: 1 RDIITTY 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_TREMBL_23:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	89.5	98	16	099VG9
2	33	86.8	315	2	052517
3	33	86.8	333	10	09MOV1
4	33	86.8	392	11	08BM6
5	33	86.8	410	10	09XE75
6	33	86.8	414	11	08BMQ7
7	33	86.8	422	10	09XE79
8	33	86.8	438	10	09XE79
9	33	86.8	753	3	08JIE3
10	33	86.8	760	3	08JIE2
11	33	86.8	2644	4	013535
12	33	84.2	39	13	090506
13	32	84.2	68	11	099LRO
14	32	84.2	179	4	096B07
15	32	84.2	242	6	095KJ3
16	32	84.2	312	4	09NZB6

17	32	84.2	343	2	08KTS8	08KTS8 candidatus
18	32	84.2	407	13	08JFE1	08JFE1 gallus gall
19	32	84.2	408	4	096EA8	096EA8 homo sapien
20	32	84.2	473	4	09BRB1	09BRB1 homo sapien
21	32	84.2	565	16	08R7J7	08R7J7 thermoaer
22	32	84.2	565	16	08R6J9	08R6J9 thermoaer
23	32	84.2	565	16	08R6J2	08R6J2 thermoaer
24	32	84.2	764	16	08PM29	08PM29 xanthomonas
25	32	84.2	816	10	08W5E9	08W5E9 oryza sativ
26	32	84.2	892	2	09LAF1	09LAF1 bacillus ce
27	31	81.6	35	16	097RE8	097RE8 streptococ
28	31	81.6	166	3	09UTC3	09UTC3 schizosacch
29	31	81.6	267	11	091W78	091W78 mus musculu
30	31	81.6	279	17	027027	027027 methanobact
31	31	81.6	307	17	028263	028263 archaeoglob
32	31	81.6	376	5	08SM17	08SM17 encephalit
33	31	81.6	447	2	048675	048675 lactococcus
34	31	81.6	543	2	093G24	093G24 streptococ
35	31	81.6	579	16	08C2D8	08C2D8 oceanobact
36	31	81.6	695	16	09A901	09A901 caulobacter
37	31	81.6	830	10	065482	065482 arabidopsis
38	31	81.6	1395	5	09W2J3	09W2J3 drosophila
39	31	81.6	1395	5	044924	044924 drosophila
40	30	78.9	168	2	051489	051489 pseudomonas
41	30	78.9	168	16	091424	091424 pseudomonas
42	30	78.9	192	17	08ZUV5	08ZUV5 pyrobaculum
43	30	78.9	281	2	092FL6	092FL6 bacillus st
44	30	78.9	297	16	092K18	092K18 rhizobium m
45	30	78.9	336	13	091669	091669 xenopus lae

ALIGNMENTS

RESULT 1

ID	099VG9	PRELIMINARY:	PRT:	98 AA.
AC	099VG9;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Hypothetical protein SAV0836 (MW0789 protein).			
GN	SAV0836 OR SA0768.1 OR SAS023 OR MM0789.			
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699),			
OS	Staphylococcus aureus (strain N315), and			
OS	Staphylococcus aureus (strain MM2).			
OC	Bacteria: Firmicutes: Bacillales; Staphylococcus.			
OX	NCBI_TaxID=158878, 158879, 196620;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Mu50, and N315;			
RX	MEDLINE=21311952; PubMed=11418146;			
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,			
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iano J.-O., Ito T.,			
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,			
RA	Makutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,			
RA	Sehmiizun K., Hiraoka H., Kohara S., Goto S., Yabuzaki J.,			
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,			
RA	Katori M., Ogasawara N., Hayashi H., Hiratake K.,			
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus			
RT	aureus";			
RT	Lancet 357:1225-1240(2001).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MM2;			
RX	MEDLINE=22040717; PubMed=12044378;			
RA	Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,			
RA	Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,			
RA	Yamamoto K., Hiratsuka K.,			
RT	"Genome and virulence determinants of high virulence community-			
RT	acquired MRSA";			
RL	Lancet 359:1819-1827(2002).			
DR	EMBL; AP003360; BAB56998.1; -.			

DR EMBL: AP003131; BAB42007.1; -
 DR EMBL: AP004824; BAB94654.1; -
 DR InterPro: IPR006663; Thiorodox_dom2.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 98 AA; 11454 MW; D33BDF1D6392C139 CRC64;

Query Match 89.5%; Score 34; DB 16; Length 98;
 Best Local Similarity 71.4%; Pred. No. 6.6;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
 ||:||||
 DB 8 RDVTFY 14

RESULT 2
 OS2517 PRELIMINARY; PRT; 315 AA.
 ID 052517;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VIRL.
 GN VIRL.
 OS Agrobacterium tumefaciens.
 OC Plasmid T1.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 NCBI_TaxID=358;
 RX MEDLINE=99009000; PubMed=9791116;
 RA Katozaki V.S., Winans S.C.;
 RT Wound-released chemical signals may elicit multiple responses from an
 RT Agrobacterium tumefaciens strain containing an octopine-type T1
 RT plasmid.
 RL J. Bacteriol. 180:5660-5667(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Winans S.C., Zhu J., Oger P.M., Schrammelfer B., Hooykaas P.J.,
 RA Fairand S.R.;
 RT Octopine-type T1 plasmid sequence."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF242881; AAC72016.1; -
 KW Plasmid.
 SQ SEQUENCE 315 AA; 36474 MW; 324B5C43B04A365 CRC64;

Query Match 86.8%; Score 33; DB 2; Length 315;
 Best Local Similarity 71.4%; Pred. No. 38;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
 ||:||||
 DB 139 RDVTFY 145

RESULT 3
 OS2517 PRELIMINARY; PRT; 333 AA.
 ID 09MOV1;
 AC 09MOV1;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE D123-1-like protein.
 GN AT4G05440.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsida.
 NCBI_TaxID=3702;
 RT [1]
 RP SEQUENCE FROM N.A.
 RA Spielgel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vill D.M.,

RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 RA Shekter M., Schutz K., See L.H., Swaby I., Haberman K., Dedhia N.N.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL161503; CAB81086.1; -
 SQ SEQUENCE 333 AA; 38777 MW; 85DBA4B4F38703D2 CRC64;

Query Match 86.8%; Score 33; DB 10; Length 333;
 Best Local Similarity 71.4%; Pred. No. 40;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
 ||:||||
 DB 198 REVTFY 204

RESULT 4
 OS8BM6 PRELIMINARY; PRT; 392 AA.
 ID 08BM6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE ADP-ribosylation factor 1 GTPase activating protein homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Plutary;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK030520; BAC27002.1; -
 SQ SEQUENCE 392 AA; 42956 MW; FC2BD8A150130CA2 CRC64;

Query Match 86.8%; Score 33; DB 11; Length 392;
 Best Local Similarity 71.4%; Pred. No. 48;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RDITTFY 7
 ||:||||
 DB 266 RDVTFY 272

RESULT 5
 OS9XE75 PRELIMINARY; PRT; 410 AA.
 ID 09XE75;
 AC 09XE75;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Patatin-like protein.
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Sorghum.
 NCBI_TaxID=4556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liaca V., Lou A., Messing J.W.;
 RT "Microarray analysis of 22-kDa zein cluster in maize and sorghum."
 DR EMBL: AF061282; AAD22170.1; -
 DR InterPro: IPR002641; Patatin.

DR Pfam: PF01734; Patatin. 1.
SQ SEQUENCE 410 AA; 44564 MW; 46405E0A445CE994 CRC64;

Query Match 86.8%; Score 33; DB 10; Length 410;
Best Local Similarity 71.4%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
:|||||
DB 95 KDLTTFY 101

RESULT 6

ID Q8BMQ7 PRELIMINARY; PRT; 414 AA.
AC Q8BMQ7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ADP-ribosylation factor 1 GTPase activating protein homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and uterus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK030295; BAC26883.1; -
SQ SEQUENCE 414 AA; 45288 MW; 72A3279D185714C0 CRC64;

Query Match 86.8%; Score 33; DB 11; Length 414;
Best Local Similarity 71.4%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
:|||||
DB 288 RDVTTF 294

RESULT 7

ID Q9XE77 PRELIMINARY; PRT; 422 AA.
AC Q9XE77;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Patatin-like protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC Liaca V., Lou A., Messing J.W.;
RT "Microsynteny analysis of 22-kDa zein cluster in maize and sorghum";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF061282; AAD22169.1; -
DR InterPro; IPR002641; Patatin.
DR Pfam: PF01734; Patatin. 1.
SQ SEQUENCE 422 AA; 45766 MW; 742128BB0A37F479 CRC64;

Query Match 86.8%; Score 33; DB 10; Length 422;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7

DB 94 KDLTTFY 100
:|||||

RESULT 8

ID Q9XE79 PRELIMINARY; PRT; 438 AA.
AC Q9XE79;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Patatin-like protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC Liaca V., Lou A., Messing J.W.;
RT "Microsynteny analysis of 22-kDa zein cluster in maize and sorghum";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF061282; AAD22149.1; -
DR InterPro; IPR002641; Patatin.
DR Pfam: PF01734; Patatin. 1.
SQ SEQUENCE 438 AA; 47495 MW; EF5EB650E2186044 CRC64;

Query Match 86.8%; Score 33; DB 10; Length 438;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RDITTFY 7
:|||||
DB 95 KDLTTFY 101

RESULT 9

ID Q8JIE3 PRELIMINARY; PRT; 753 AA.
AC Q8JIE3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cellulase Cel48A precursor.
RN [1]
OS Cel48A.
OS Plomycetes sp. E2.
OC Eukaryota; Fungi; Chytridiomycota; Neocallimasticales;
OC Neocallimasticeae; Plomycetes.
OX NCBI_TaxID=73868;
RN [1]
RP SEQUENCE FROM N.A.
RC Steenbakkers P.J.M., Freelove A., van Cranenbroek B., Sweegers B.M.C.,
RA Harhangi H.R., Vogels G.D., Hazlewood G.P., Gilbert H.J.,
RA Op den Camp H.J.W.;
RT "The Major Component of the Cellulosomes of Anaerobic Fungi from the
RT Genus Plomycetes is a Family 48 Glycoside Hydrolase."
RL DNA Seq. 13:313-320(2002).
DR EMBL; AF449412; AAN76734.1; -
KW SIGNAL.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN. 25 753 CELLULASE CEL48A.
SQ SEQUENCE 753 AA; 83489 MW; D7C28317A4AB7BBF CRC64;

Query Match 86.8%; Score 33; DB 3; Length 753;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DITTFY 7
:|||||
DB 414 DITTFY 419

RESULT 10

08J1E2
ID 08J1E2 PRELIMINARY; PRT; 760 AA.
AC 08J1E2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cellulase Cel48a precursor.
GN CEL48A.
OS Pteromyces equi.
OC Eukaryota; Fungi; Chytridiomycota; Neocallimastixales;
OC Neocallimastixaceae; Pteromyces.
OX NCBI_TaxID=99929;
RN [1]
RP SEQUENCE FROM N.A.
RA Steenhakkers P.J.M., Freejove A., van Craenenbroek B., Sweegers B.M.C.,
RA Harhangi H.R., Vogels G.D., Hazlewood G.P., Gilbert H.T.,
RA Op den Camp H.J.M.;
RT "The Major Component of the Cellulosomes of Anaerobic Fungi from the
RT Genus Pteromyces is a Family 48 Glycoside Hydrolase."
RL DNA Seq. 133313-320(2002).
DR EMBL; AF449413; AAN76735.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 760 CELLULASE CEL48A.
SQ SEQUENCE 760 AA; 84648 MW; B7AF925F95EB0271 CRC64;
Query Match 86.8%; Score 33; DB 3; Length 760;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DIFFY 7
DB 415 DIFFY 420

RESULT 11
013535
ID 013535 PRELIMINARY; PRT; 2644 AA.
AC 013535; Q93051;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE FRAP-related protein (Protein kinase ATR).
GN ATR OR FRPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96181495; PubMed=8610130;
RA Clamprich K.A., Shio T.B., Keith C.T., Schreiber S.L.;
RT "cDNA cloning and gene mapping of a candidate human cell cycle
RT checkpoint protein."
RL Proc. Natl. Acad. Sci. U.S.A. 93:2850-2855(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Bentley N.J., Holtzman D.A., Flagg G., Keegan K.S., Demaggio A.,
RA Ford J.C., Hoekstra M., Carr A.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49844; AAC50405.1; -.
DR EMBL; Y09077; CAA70298.1; -.
DR EMBL; U76308; AAC50929.1; -.
DR Genbank; HGNC:882; ATR.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000357; HEAT_repeat.
DR InterPro; IPR000403; PI3_Pi4_kinase.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_Pi4_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PSS0077; HEAT_REPEAT; 1.

DR PROSITE; PSS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PSS0290; PI3_4_KINASE_3; 1.
KW Kinase; Transferase.
FT CONFLICT 92 92 R -> A (IN REF. 1).
SQ SEQUENCE 2644 AA; 301449 MW; 3889CAB930E630A CRC64;
Query Match 86.8%; Score 33; DB 4; Length 2644;
Best Local Similarity 85.7%; Pred. No. 3.e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DIFFY 7
DB 1814 RDITAFY 1820

RESULT 12
090506
ID 090506 PRELIMINARY; PRT; 39 AA.
AC 090506;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Initiation factor (Fragment).
GN EIF-4A.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary;
RA Price D.A., Lin Y.-W.P.;
RT "PCR fragment from Killifish that is homologous to mouse EIF-4A."
RL EMBL; M89792; AAA49287.1; -.
DR NON_TER 1 1
SQ SEQUENCE 39 AA; 4530 MW; 824F368975B22686 CRC64;
Query Match 84.2%; Score 32; DB 13; Length 39;
Best Local Similarity 85.7%; Pred. No. 7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DIFFY 7
DB 18 RDITAFY 24

RESULT 13
099LR0
ID 099LR0 PRELIMINARY; PRT; 68 AA.
AC 099LR0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 7.9 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002266; AAH02266.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 68 AA; 7882 MW; BDD7476AB5499E88 CRC64;
Query Match 84.2%; Score 32; DB 11; Length 68;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITFFY 7
 ||| |||
 DB 47 RDITFFY 53

RESULT 14

Q96B07 PRELIMINARY; PRT; 179 AA.
 AC Q96B07;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DDbj databases.
 DR EMBL: BC016295; AAH16295.1; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00490; HELIC; 1.
 KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 179 AA: 21039 MW: 325A790227A89C35 CRC64;

Query Match 84.2%; Score 32; DB 4; Length 179;
 Best Local Similarity 85.7%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITFFY 7
 ||| |||
 DB 158 RDITFFY 164

RESULT 15

Q95KJ3 PRELIMINARY; PRT; 242 AA.

AC Q95KJ3;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
 DE Hypothetical 28.5 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_Taxid=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal cortex;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT *Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries.*;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DDbj databases.
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 DR EMBL: AB060830; BAB46863.1; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD box.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00490; HELIC; 1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 242 AA: 28492 MW: 03D02D3A7B405A21 CRC64;

Query Match 84.2%; Score 32; DB 6; Length 242;
 Best Local Similarity 85.7%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RDITFFY 7
 ||| |||
 DB 221 RDITFFY 227

Search completed: October 7, 2003, 19:20:05
 Job time : 8.97586 secs

